

Edith Cowan University
Research Online

Theses : Honours

Theses

1989

The evaluation of writing in unit curriculum English

David Sharp
Edith Cowan University

Follow this and additional works at: https://ro.ecu.edu.au/theses_hons



Part of the [Curriculum and Instruction Commons](#)

Recommended Citation

Sharp, D. (1989). *The evaluation of writing in unit curriculum English*. https://ro.ecu.edu.au/theses_hons/171

This Thesis is posted at Research Online.
https://ro.ecu.edu.au/theses_hons/171

Edith Cowan University

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement.
- A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

THE EVALUATION OF WRITING IN UNIT CURRICULUM ENGLISH

David Sharp BEd

School of Education 7th July 1989

(II)

ABSTRACT

The assessment and grading of writing in Unit Curriculum English plays a major role in the determination of a student's summative, and hence, public, letter grade. Through teachers' adherence to the assessment and grading procedures for writing in Unit Curriculum English, the Ministry of Education lay claims on comparability and statewide standards. The claim warranted the investigation of the guidelines and procedures used.

A review of literature on the evaluation of writing was conducted. In order that local application and relevance be possible, the holistic mode of evaluating writing was focused on. Problems were identified in the research concerning score reliability. The pre-requisites for obtaining statistically reliable scores were outlined. The pre-requisites include training and monitoring scorers to apply the established evaluative criteria to pieces of writing. The research highlighted that, in spite of extensive training and monitoring, problems of reliability remained. This was attributed to the fact that scorers cannot always adhere to the evaluative criteria specified in holistic grading procedures. It was pointed out that scorers' conceptualisations of writing proficiency differ. The face validity of the evaluative criteria were therefore subject to disagreement.

(III)

These findings were discussed in relation to the assessment and grading procedures for writing in Unit Curriculum English. It was pointed out that as many teachers of English are inexperienced and untrained in holistic evaluative procedures, the validity of the evaluative criteria for writing in Unit Curriculum English were open to question. This exacerbated problems of the reliability of grades awarded under Unit Curriculum English. In the light of these findings, the credibility of the Ministry of Education's claims on comparability and statewide standards in Unit Curriculum English were questioned.

A conceptualisation of writing proficiency in Unit Curriculum English was offered. The conceptualisation highlighted the product emphases for writing in Unit Curriculum English. Determining the degree to which the current assessment and grading procedures addressed these emphases highlighted problems and shortcomings. The findings supported the research by identifying a number of factors which placed the reliability of grades in Unit Curriculum English at risk.

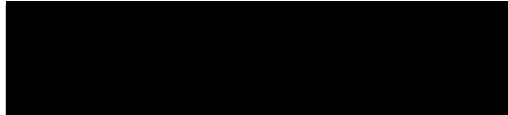
The paper establishes that the current guidelines and procedures for assessing and grading writing in Unit Curriculum English are lacking as they fail to adequately address the pre-conditions of reliable scoring.

(IV)

Concomitantly, the credibility of the Ministry of Education is at risk. In order to achieve comparability and statewide standards, reliable scoring must occur. To redress the risk, the insufficiencies of the guidelines and procedures, to which teachers of Unit Curriculum English comply, need to be addressed. Considering the political and ideological dimensions of education policy, it was felt that failure to redress these insufficiencies would reflect more poorly upon teachers of Unit Curriculum English than it would the Ministry of Education.

(V)

I certify that this thesis does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any institution of higher education and that, to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where due reference is made in the text.



(VI)

ACKNOWLEDGEMENTS

I am indebted to my supervisor, Mr Ken Willis, for his patience, professionalism and objective and honest advice. Without his guidance, this paper would not have been possible. I am also indebted to Mrs Carol Hogan for her critical remarks at a crucial time. And last, but no means least, thanks go to my parents for their tireless support and interest.

CONTENTS

Title Page		
Abstract	II	
Declaration	V	
Acknowledgements	VI	
Contents	VII	
Introduction	1	
Section One		
Unit Curriculum English		7
Section Two		
The Evaluation of Writing		14
Methods of Determining Writing Proficiency		17
The Holistic Evaluation of Writing		22
Analytic Scales		23
General Impression Marking		25
Essay Scales		25
Reliability and Validity		26
Discussion		35
Recommendations		37
Section Three		
A Conceptualisation of Writing Proficiency		39
Analytic Scales		50
Recommendations		53
Section Four		
Improving Assessment and Grading		55
Closing Statement		64
List of Recommendations		65
Appendices		
Appendix A		69
Appendix B		73
Bibliography		79

INTRODUCTION

Today, conceptualisations of the terms, credibility and accountability, underpin the operations of service systems and institutions. The terms encapsulate the essence of political rhetoric which has a marked influence over public policy development and action, and education, rightly or wrongly, has not been immune from their ideological or political implications. The terms appear regularly in education policy statements and, while one may be accustomed to expect their application to matters of an administrative nature, one increasingly finds them being extended to apply to areas of teaching personnel and practice. In Western Australia, the terms' rise to prominence accompanied a shift in emphasis from process to product based schooling, sparked by the need for schools to become more responsive to societal and technological change and demand.

The link between the perceived needs of the public and education policy is important, as they share a continuum of being. That is, one shapes the other. (For useful sources on the ideological dimensions of educational policy, refer to Johnston, 1983; Marginson, 1985; McKinnon, 1982; Taylor, 1982; Wirt, 1979.) More important, however, is the fact that education policy influences work-face operations at school level. In a time of economic belt-tightening and increased public criticism, it is not surprising that the trends are now

as prevalent as they are, and extended into newer, more defined areas of applicability such as those mentioned above. As Popham (1981) points out, there is no doubt we are living in the middle of an "evidence-oriented era".

Popham's view aptly applies to education and schooling. The evidence-oriented phenomenon has filtered down the education hierarchy to touch those directly concerned with teaching practice and, in an attempt to become more credible and accountable, schools and teachers have had no choice but to become evaluation conscious.

This is not to suggest that evaluation is a recent innovation of schooling. A suggestion such as this would be artless to say the least. It merely refers to the fact that in the current climate, evaluation has found both elevated status and increased utilisation in schools and teaching practice.

This is particularly true of the Western Australian scene. The perceived need for schools and education to become more responsive to societal and technological change and demand - more credible and accountable - prompted judgements that, in 1988, saw the abolition of the lower secondary school Achievement Certificate

system (Dettman 1969) and the implementation of the Unit Curriculum (Beazley 1984). Sparked by policy initiatives of the evidence-oriented era, the Unit Curriculum represents this State's most ambitious attempt to address the trends of credibility and accountability in schooling. The effects of the change-over are still being felt by secondary school teachers and administrators. While the Unit Curriculum has brought a new face to secondary schooling in Western Australia, it has also brought about changes to teaching procedure and practice.

It should be of little surprise to learn, then, that great effort has been made by the Ministry of Education to clarify the operational procedures for teachers of the Unit Curriculum. It will be of no surprise that particular attention within these new procedures has been given to the monitoring and evaluation of student performance.

This paper is concerned with the evaluation of student performance. More specifically, an investigation will be made of the Ministry of Education's guidelines and procedures for evaluating writing within Unit Curriculum English.

This paper adopts a format which begins with a broad description of Unit Curriculum English. This is to provide the reader with background information and is in no sense analytical. The section entails an historical perspective to the syllabus and discusses its relationship to the recommendations of the 1984 Committee of Inquiry into Education in Western Australia (Education in Western Australia), chaired by Mr Kim Beazley (subsequently referred to as the Beazley Report). Some brief descriptive notes follow in regard to unit structure and composition before proceeding to an overview of the syllabus' evaluation guidelines. Here, the paper will identify the type of evaluation used by the Ministry of Education with respect to writing in Unit Curriculum English, namely, what it refers to as holistic grading.

The second section of the paper presents a review of significant research and theory pertaining to the evaluation of writing. In order that local application and relevance be possible, the holistic mode of evaluating writing is focused on. This review of literature concludes with broad suggestions for further research and identifies two key problems facing education authorities and teachers of English.

The third section of the paper offers a conceptualisation of writing proficiency based on content analysis of the English Unit Curriculum documents. The conceptualisation will be discussed in terms of the research and theory presented in the preceding section.

The fourth section concerns itself primarily with the guidelines and procedures for evaluating writing as found in the Unit Curriculum English documents. It aims to identify any insufficiencies in clarity and explicitness. In this sense, the section aims to prompt improvements to the guidelines and procedures so as to benefit those for whom they were designed - the teachers of English. A closing statement and full list of recommendations will end the paper.

SECTION ONE

Unit Curriculum English

The Committee believes that the community expects higher standards from more of the graduating students. It is assumed that basic competencies are a fundamental right of as many students as possible and that a greater number than ever before should be able to communicate effectively and accurately, (and) understand what they read. The education system must address itself seriously to this ideal.

The Committee of Inquiry into Education in Western Australia.

The Unit Curriculum English syllabus was developed in light of recommendations of the Beazley Report of 1984. In reviewing the then existing syllabus, the Committee of Inquiry identified three major points of concern. The first was that the existing syllabus for English was inadequate in that it lacked prescription (p.144). The chief concern centred around the possibility of teachers not addressing important issues not explicitly stated within the syllabus, and particular reference was made to literacy skills (p.147). The second point of concern related specifically to literacy and "functional English" (p.125). At the time a very public issue, literacy was marked as an area requiring "comprehensive" attention (p.29). Indeed, the Committee went as far as to define "literate" and itemise specific skills that constitute "functional" or "literate" individuals (p.123-125). The third point of concern related to evaluation, and called for more emphasis on formal and informal testing within the process and practice of English teaching, so as to allow for the monitoring and evaluation of students' literacy skills (p.147).

These three concerns assume significant positions within Unit Curriculum English, and are clearly echoed in what amounts to a detailed series of syllabus documents that place value on the acquisition, development, monitoring and evaluation of functional communicative skills.

Obviously enough, these concerns are explicated within the "process" and "text" objectives of each unit and reinforced by the suggested activities which accompany them. The Unit Curriculum's English syllabus offers twenty two units spread across six stages of progress. Each unit falls into one of three categories: Focus, General, or Special Emphasis.

As their names suggest, each category offers units designed to cater for varying ability levels across the three years of compulsory secondary schooling. The Focus units (Stages 1 and 3) have been tailored to cater for those students with poorly developed language skills, while the Special Emphasis units (Stages 4 to 6) cater for the more able students with particular abilities, interests and needs. The large majority of students will study units from the General category (Stages 2 to 6). These units have been designed to provide students with a sound language skills schooling experience, and cater to the syllabus' concern for the development of competence in literacy skills.

Pathways (a term used to describe students' progress through the unit continuum), are normally selected on the basis of the students' results in the first year of secondary schooling. Pathways are ideally chosen to

match students' abilities and are both varied and flexible. A student of average ability for instance, may enter the continuum at Stage 2 and exit, after three years, at Stage 6. Alternatively, the same student may proceed no further than Stage 5 and, in doing so, would be required to study more units at the chosen exit level. A student of exceptional ability may enter at Stage 3 and exit at Stage 6, with the majority of units studied coming from the Special Emphasis category, while a student of lower than average ability would enter at Stage 1 and exit at Stage 4. Table 1 presents typical pathways for each ability level.

<u>BELOW AVERAGE ABILITY</u>						
STAGE	1	2	3	4	5	6
UNIT	1.1#	2.1	3.1	4.1	5.1	6.1
	1.2#	2.2	3.2	4.2	5.2	6.2
			3.3#	4.3*	5.3*	6.3*
			3.4#	4.4*	5.4*	6.4*
						6.5*
						6.6*
<u>AVERAGE ABILITY</u>						
STAGE	1	2	3	4	5	6
UNIT	1.1#	2.1	3.1	4.1	5.1	6.1
	1.2#	2.2	3.2	4.2	5.2	6.2
			3.3#	4.3*	5.3*	6.3*
			3.4#	4.4*	5.4*	6.4*
						6.5*
						6.6*
<u>ABOVE AVERAGE ABILITY</u>						
STAGE	1	2	3	4	5	6
UNIT	1.1#	2.1	3.1	4.1	5.1	6.1
	1.2#	2.2	3.2	4.2	5.2	6.2
			3.3#	4.3*	5.3*	6.3*
			3.4#	4.4*	5.4*	6.4*
						6.5*
						6.6*

KEY

Focus Unit
 * Special Emphasis Unit
 Green Year 8
 Orange Year 9
 Red Year 10

Table 1: Typical Pathways across ability levels in Unit Curriculum English

Students will normally complete four units in a year. The minimum study requirements and availability of units (with the exception of those from the General category, of which all, or most, must be offered) will, in part, determine the pathways of progress through which students proceed. These matters, as well as those concerning recording and reporting procedures, are issues left for schools to resolve.

As stated earlier, each English unit consists of both process and text objectives. The process objectives are common to all units within the English syllabus and are shown below in Table 2.

Process Objectives

All Units aim to develop students' abilities to:

1. use the conventions of standard English in writing;
2. prepare and participate in a range of one to one, individual and group oral language activities;
3. understand, order and convey facts, ideas and opinions in a variety of comprehending and composing situations;
4. understand and respond to structure, style and tone and vary language according to audience and purpose; and
5. understand and use a wide vocabulary.

Table 2. Process Objectives of the Unit Curriculum
English Syllabus

The generality of these objectives allows for teachers of English to integrate any number of these with a unit's text objectives which are task specific and fall under the headings of: transactional, media, prose fiction, poetry and drama. Accompanying these text objectives are suggested activities which serve as guides to the type of work students are to submit for purposes of evaluation.

In terms of evaluation structures, the process objectives are expected to account for approximately 70 per cent of a student's grade in any unit, while the remaining 30 per cent is allocated toward text objectives. Of this 30 per cent, suggested weightings for the text headings are provided with each unit description and, with the exception of Stage 1 Focus units, writing and reading weightings account for 60-70 per cent of the total allocated. As to what proportion of the 60-70 per cent allotment constitutes writing or reading is unclear, and would presumably be determined at school departmental level.

The evaluation structures within units are designed to be integrated with the Unit Curriculum's standards-referenced assessment and grading procedures.

Previously, students were awarded grades based upon norm-referenced procedures, that is, they were awarded

grades relative to the performance of others. With the introduction of standards-referencing, pre-determined standards are used as templates during the evaluation process. Teachers match the work of their students to these standards to arrive at a representative letter grade that signifies a level of performance. In most instances, it is expected that grades be awarded using an holistic approach.

It may pay at this stage to highlight the point that when viewed holistically, the assessment weighting allocated towards writing in Unit Curriculum English is considerable. Writing not only accounts for some portion of the 60-70 per cent text objective allotment, but also for a large portion of the process objective allotment. The most tangible means to assess process objective attainment (with the exception of the second process objective), is through the evaluation of students' writing. One would anticipate then, that the consistent or reliable evaluation of writing underpins the Ministry of Education's claims on comparability and statewide standards in Unit Curriculum English. When we combine this point with that concerning the marking loads confronting teachers of Unit Curriculum English, we arrive at an issue worthy of serious investigation.

While the evaluation of written work is a conventional part of English teaching practice, we have yet to determine the full extent of the ramifications of such a convention in the light of Unit Curriculum time-frames and procedures. The evaluation of writing in Unit Curriculum English, then, assumes great significance and raises certain questions. For instance, can statewide standards be consistently maintained when evaluating such a complex process as writing? Are teachers over-evaluating in Unit Curriculum English? How useful are the procedures documents for assessing and grading writing?

Questions of this type highlight three important issues. Firstly, that an investigation of the guidelines and procedures for evaluating writing in Unit Curriculum English is warranted. Secondly, that issues raised in the investigation prompt more in-depth research, and thirdly, that the findings of such research be directed to benefit English teachers and English teaching practice.

The next section of this paper will deal with the evaluation of writing. It aims to highlight significant research and theory that bears impact on the assessment and grading procedures of Unit Curriculum English. Discussion will also be made of the possible effects these procedures have on the Ministry of Education's claim on comparability and statewide standards in Unit Curriculum English.

SECTION TWO

The Evaluation of Writing

My predominant impression has been that writing classes are fantastically over-evaluated. Students are graded on everything they do every time they turn around. Grades generate anxiety and hard feelings between everyone. Common sense suggests that grades ought to be reduced to the smallest possible number necessary to find out how students are getting along toward the four or five main objectives of the program, but teachers keep piling them up like squirrels gathering nuts...

Paul Diederich

In the preceding section of this paper the point was made that all facets of education and schooling had become evaluation conscious. Such a phenomenon was tied to what Popham (1981) refers to as the "evidence-oriented era" in which we live. It was pointed out that through the implementation of education policy directives, this phenomenon was affecting teaching practice in Western Australian secondary schools. This view was substantiated by the fact that, within the realm of teaching Unit Curriculum English, the Ministry of Education had produced a series of syllabus documents that made explicit the requirements of monitoring and evaluating literacy skills. It was concluded that such skills - through virtue of the syllabus' process and text objectives, its assessment structures and guidelines as to weightings - were best measured through the evaluation of students' writing. While acknowledging the fact that literacy refers to all aspects of the communicative process, it was pointed out that writing provides the most visible, and therefore most tangible medium from which to make judgements concerning the acquisition and development of language competencies. It was concluded further that the evaluation of writing assumed an integral component of Unit Curriculum English assessment, and that reliable evaluation underpins not merely the facilitation of teachers' informed educational decisions, but also the Ministry of Education's claims on statewide

comparability of grades awarded under Unit Curriculum English.

In its procedures for the assessment and grading of writing in Unit Curriculum English, the Ministry of Education states that students' grades are to be awarded according to pre-determined standards as provided within the syllabus documents, via a holistic evaluative procedure. The holistic mode of evaluating writing therefore deserves discussion in the light of significant research findings and related theory.

Before proceeding any further, it is necessary to provide definitions of the terms which will appear regularly in the review. This is to serve reasons of clarity and brevity. The definitions, excluding the author's, have their roots in the descriptive dialogue of educational research and are generally accepted as being accurate in description.

Evaluation is central to the theme of this paper, and when so used refers to Gay's definition. Gay (1985) defines evaluation as the systematic process of collecting and analysing data in order to make decisions. Its purpose is to determine the status of the object of evaluation and to compare it to a set of standards or criteria. (p.370)

Evaluation in this paper is concerned with writing. For this paper, **writing** refers to the range of composed written discourses of students submitted as required to teachers of Unit Curriculum English for purposes of evaluation.

Two further terms which are important to this paper are validity and reliability. **Validity** refers to the degree to which a test measures what it is supposed to measure, while **reliability** refers to the degree of consistency of scores of a given test (Biggs & Telfer 1981). If scores of a given test can be replicated it is said to be reliable.

Methods of Determining Writing Proficiency

There are two methods of determining writing proficiency. These methods are characterised as being either indirect or direct. The indirect method, also referred to as the quantitative or objective method, involves the use of standardised tests. These tests assess students' ability to identify differences between standard and non-standard English usage. Stiggins (1982) points out that students respond to "a series of objective test items which often follow a multiple choice format, (in which) actual writing is not required" (p.348). Charney (1984) identifies an occasionally used

variation of this method that combines the standardised test with a written sample. Here, the method retains its indirect label as scores for written samples are determined through the counting of grammatical errors, the number of t-units (syntactical sophistication), or uncommon vocabulary items (p.66). In either case, both test students' knowledge of the surface conventions of effective writing, such as grammar, diction, punctuation, spelling, and sentence order and construction. (Charney, 1984; Cooper & Odell, 1977; Stiggins, 1982.)

The direct or qualitative method of determining writing proficiency requires students to compose a piece of written discourse for evaluation. As the description implies, this method places emphasis on composition. Marks or grades are awarded according to the degree to which the writing satisfies certain standards or criteria that cannot be assessed using indirect methods, such as flavour, impact, purpose, and argument.

Each of these methods has its relative merits and pitfalls and it may prove prudent at this stage to identify them. As already described, the indirect method of assessing writing proficiency requires the use of standardised tests. Answers are made in multiple choice

format and scored either manually or by machine. In the case of an essay accompanying the test, tallies are made of the essay's surface attributes and usage errors. The first advantage of this method then, concerns its ease of scoring which is relatively inexpensive, and less time-consuming than that of the direct method (Stiggins 1982). Secondly, scores obtained using the indirect method will yield a very high degree of statistical reliability. (Charney, 1984; Cooper & Odell, 1977; Culpepper & Ramsdell, 1982; Lloyd-Jones, 1977; McColly, 1970; Veal & Hudson, 1983.) Thirdly, the indirect method empowers the testing agent (or body) to focus on specific skills by virtue of selecting appropriate test items. As Stiggins (1982) points out, there is a high "degree of control over the nature of skills tested." (p.356) Finally, if standardised tests are criterion referenced, Stiggins claims they can serve diagnostic purposes (p.356). This search of literature failed to ascertain whether such a claim had been tested. However, in view of the sophisticated computer software now available, it seems Stiggins' claim would not be untenable. Such use of the indirect method would prove useful to language researchers.

The major advantage of the direct method of determining writing proficiency lies with the fact that scorers are engaged with a student's written composition during the process of evaluation. Thus, more information about writing proficiency is provided to scorers using direct methods than to those using indirect methods. This relative wealth of evidence appeals to language theorists such as Cooper and Odell (1977), who claim direct methods are a more valid means to evaluate writing proficiency. Similar views underline the work of Britton, Martin and Rosen (1966), Diederich, French and Carlton (1966), Gere (1980), Halliday (1978), Hirsch (1977), and Lloyd-Jones (1977) . Because of the evidence available to scorers, direct methods are useful for diagnostic purposes and, of particular interest to teachers, can serve as instructional aids in the teaching of writing (Stiggins, 1982).

Just as there are advantages with each method, so too are there disadvantages. As one would expect with objective test formats, indirect methods testing can be as much a test of reading comprehension as they can be of writing proficiency. It is this point opponents of the method use in support of their argument that indirect methods lack validity. Cooper and Odell (1977) point out further, that as tests of this type assess the editorial

skills of writing, and do not require students to compose written discourse, they lack the real-world application which direct testing methods satisfy. Again, similar concerns are reflected and the works of Britton et al. (1966), Gere (1980), Hirsch (1977), Liner and Kirby (1981), Lloyd-Jones (1977), Walshe, March and Jensen (1986), and Wilkinson (1980), who are all proponents of the direct method.

Direct methods have two traditional disadvantages. Stiggins (1982) identifies the first as costliness, which stems from the time involvement associated with training scorers and marking papers. Marking and grading papers requires scorers to appraise each extended composition for its intrinsic qualities and, as such, requires more time commitment to the evaluation process. This problem is exacerbated when dealing with large-scale marking situations. The second traditional disadvantage of the direct method concerns reliability. This will be discussed in detail later. Another problem which will also be discussed later, reflects more recent research findings and concerns the validity of direct methods assessment.

It is worthy to note that in Western Australia, indirect methods of determining writing proficiency are not used in the mainstream evaluation procedures of lower secondary school English. It would be remiss of this investigation, nevertheless, to fail to acknowledge a possible use for it. This will be discussed towards the end of this section.

Having now described the two alternative methods of determining writing proficiency and their advantages and disadvantages, attention will now be given to the holistic mode of evaluating writing. Holistic evaluation is a form of direct methods assessment and, as previously stated, is the mode of evaluation specified for use in the Unit Curriculum guidelines for assessing and grading writing in English.

The Holistic Evaluation of Writing

Holistic rating is a quick, impressionistic qualitative procedure for sorting or ranking samples of writing. It is not designed to correct or edit a piece, or to diagnose weaknesses. Instead, it is a set of procedures for assigning a value to a writing sample according to previously established criteria.

(Charney, 1984, p.67)

Holistic grading procedures were devised as a means to overcome the problems of reliability associated with the direct method of scoring and grading essays. While it was generally accepted that direct methods were the most valid means to evaluate writing, research identified that scores among markers were inconsistent. Not only did different markers award different grades to the same paper, but single markers tended to award different grades to the same paper at different times. (Wesdorp et al. 1982.) As direct grading procedures gained widespread acceptance, research concentrated on finding ways to improve existing methods. The result is a number of holistic evaluation types, all of which have yielded reliable scores given certain conditions. Some of these types are described briefly below.

Analytic Scales

An analytic scale is comprised of a list of features common to a writing mode. Each feature is categorised under a general trait and divided into three levels: low, middle, and high, to which numerical values are attached. The numerical values of each feature vary according to their importance within the mode of discourse being evaluated. In order that scorers become attuned to each feature's value, descriptive notes provide cues as to what distinguishes low, middle and high level responses.

	Low		Middle		High
<u>General Merit</u>					
Ideas	2	4	6	8	10
Organisation	2	4	6	8	10
Wording	1	2	3	4	5
Flavour	1	2	3	4	5
<u>Mechanics</u>					
Usage	1	2	3	4	5
Punctuation	1	2	3	4	5
Spelling	1	2	3	4	5
Handwriting	1	2	3	4	5
				Total	

Table 3. Analytic Scale Score Sheet for the Evaluation of Writing (Diederich, 1974).

As Cooper (1977) points out, the double-weighting given to the "ideas" and "organisation" features of this scale, reflect the main points of emphasis during evaluation, and are deemed to be important to the mode of discourse being evaluated. Analytic scales thus have a fair degree of adaptability while providing an explicit list of important and distinguishable features. Analytic scales are attributed to the research of Diederich et al. (1966) and have since become more detailed and sophisticated in design. Cooper (1977), states that due to their explicit nature, the more recently developed analytic scales can serve useful diagnostic ends. For two such examples, refer to the Appendix section of this paper.

General Impression Marking

General impression marking requires the scorer to decide where a paper fits in relation to others. No scoring guides are needed. Cooper (1977) describes that the scorer simply ranks each paper to arrive at an ordered series ranging from best to worst, based upon implicit notions of what characterises a good paper from a bad one. This method has produced reliable results. Cooper (1977) points out that Britton et al. obtained reliabilities as high as .82. General impression marking is commonly used in tertiary institutions both within and outside of Australia, generally to good effect.

Essay Scales

Essay scales are an ordered series of complete essays arranged according to writing quality. They provide a full range of the types of level of response (eg. A to F) and are often accompanied by brief summary statements on the attributes of each sample. A scorer attempts to match a piece of writing to those provided in the scale to arrive at a score or grade. Researchers prominent in this area include Nail et al. (1960), who developed a scale for expository essays, and Martin et al. (1965) who developed a scale for imaginative writing. Essay scales,

labelled "exemplars" in Western Australia, are integral to the Ministry of Education's standards-referencing evaluation system in Unit Curriculum English.

Absent from this description of holistic evaluation types are those which have limited application to the grading of students' writing in Unit Curriculum English.

Dichotomous scales and the Centre of Gravity response marking schemes are two examples. (Both these can serve useful formative roles, however.) Absent also are those procedures which would prove impractical for school usage. They include Primary Trait Scoring (Lloyd-Jones 1977) and the Wilkinson scales (Wilkinson 1980). For information concerning these evaluation types refer to Cooper and Odell (1977) and Verhulst (1987).

Reliability and Validity

Research related to the holistic evaluation of writing has long identified problems concerning the reliability of scores. (Wesdorp et al. 1982.) In an attempt to ascertain the causes of unreliability, Diederich, French and Carlton (1961) attempted to determine whether schools of thought existed among scorers of essays. Using sixty scorers from six backgrounds, scorers were asked to mark each of the three hundred papers provided. The results

were analysed and indicated that schools of thought did exist among scorers. The correlations between schools of thought were insignificant but reliability of scores within schools of thought were significant. The conclusion was made that for holistic scores to yield reliable figures, scorers should come from similar backgrounds. (McColly, 1970; Wesdorp et al., 1982.)

Research also identified that there are variables within essays that influence holistic scores, irrespective of whether scorers have similar backgrounds and training in holistic methods. Freedman (1979) confirmed the research of Harris (1977) when she found that scores given to essays with re-written or manipulated components of organisation and content, correlated to the degree of "manipulation".

Her findings confirmed that organisation and content were "powerful" influences on holistic scores (p.337). She posited the notion of a "hierarchy of values" where, given that essays were well organised, sentence structure and mechanics increasingly became more influential. Thus, an essay strong in organisation and content would have a score significantly influenced by sentence structure and mechanical prowess. The research of Breland and Jones (1984) confirmed this finding.

In 1979, Stewart and Grobe conducted research that showed that scorers were more influenced by essay length and freedom from surface usage errors than they were by the "syntactic resources of language" shown in students' expository essays (p.75). Grobe (1981) replicated this study to see whether similar outcomes would apply to the scoring of narrative essays. His results indicated that narrative essays were also scored significantly higher if they were lengthy and free of surface errors. He concluded that what is "perceived as good narrative writing is closely associated with vocabulary diversity" (p.85). As dexterous use of vocabulary would distract scorers from the other, more important aspects of writing, Grobe warned that to "state that schools should concentrate on improving childrens' vocabulary in order to improve their writing... would most likely be a mistake" (p.85). This conclusion bears significance to the fifth process objective of the Unit Curriculum English syllabus (refer Table 2). The research findings of Breland and Jones (1984) also supported those of Stewart and Grobe (1979) and Grobe (1981), in that they too, found essay length, freedom from surface usage errors and vocabulary to be significant influences on holistic scores.

Using a similar approach to the Freedman (1979) study, Hake and Williams (1981) conducted research to ascertain the degree to which style and vocabulary affected holistic scores. Their findings showed that essays re-written in nominal style, as opposed to those re-written in verbal style, received higher scores by virtue of so-called "superior logic and organisation". This is in spite of the fact that both essay styles were constructed to be identically organised, argued and supported (p.437). They concluded that a dexterous vocabulary implied intellectual maturity (p.440), and that such written essays would receive higher scores than pieces written with a less developed vocabulary, in spite of content. Neilsen and Piche (1981) also found vocabulary to be a significant influencing factor on holistic scores regardless of syntactical complexity (p.71).

Daly and Dickson-Markman (1982) investigated the degree to which "context effects" (the influence of previous stimuli on subsequently presented stimuli) influenced holistic scores. They found that when "an average essay is read after a series of high quality pieces, it is rated lower than when it is preceded by a group of low quality ones." (p.313) Similarly, Hughes, Keeling and Tuck (1983a) found that context effects persisted even

when scorers were instructed upon how to guard against them. (p.1049) Of particular relevance to the local scene was a subsequent study of Hughes and Keeling (1984) which aimed to determine whether model essays, such as those found in the essay scales of the Unit Curriculum English documents, reduced context effects. It was found that model essays had no effect whatsoever. (p.280)

Finally, Breland and Jones (1984) found that handwriting quality and neatness were influencing factors on holistic scores. (Refer also McColly, 1970.) Their finding supported that of Hughes, Keeling and Tuck (1983b), and Markham (1976).

The claim that holistic evaluation procedures are a valid means to determine writing proficiency rests with the assumption that direct methods assessment allows for the evaluation of "real" and higher-order writing skills. As the research findings illustrate, and somewhat paradoxically it might be added, scorers of essays using holistic methods are significantly influenced by the more mundane, surface level characteristics of effective writing. Essay length, freedom from usage errors, vocabulary, spelling and handwriting neatness and appearance consistently influence holistic essay scores.

Pre-requisites for the reliable holistic scoring of essays, says Cooper (1977), are dependent upon the scorers' coming from similar backgrounds and being carefully chosen, trained and monitored while scoring (p.18). This claim is supported by McColly (1970), Sweedler-Brown (1985), and Wesdorp et al. (1982). These pre-requisites raise a point of concern, and reflect more recent research and theory regarding the validity of holistic scoring.

The concern centres around the degree to which scorers are made reliable. Charney (1984) points out that training procedures are designed to sensitise readers to the agreed criteria, and guide them to employ those standards rather than their own. She states that three methods are used during training and scoring sessions to ensure "complyability". The first is peer pressure, and she cites the research findings of Coffman;

In general, when made aware of discrepancies, teachers tend to move their own ratings in the direction of the average ratings of the group. Over a period of time, the ratings of staff as a group tend to become more reliable. (p.74)

The second method of ensuring "complyability" is monitoring.

Monitoring by "table leaders" is also a common practice. It is useful for detecting variance, caused in some cases by the onset of fatigue in the readers, which would reduce the statistical reliability of the results. (p.74)

The third point is rating speed. Here, Charney refers to the work of McColly (1970).

If a reader is competent, and if he has been well-trained and oriented, his instantaneous judgement is likely to be a genuine response to the thing for which he is looking. But if he is given time to deliberate, he is likely to accommodate his judgement to tangential or irrelevant qualities which will introduce bias into the judgement.

(p.74)

Charney concludes that, "it seems that in order to achieve high reliability, testing agencies and researchers must impose a very unnatural reading environment, one which intentionally disallows thoughtful responses to essays." (p.74)

These points are presented in detail to highlight the fact that for holistic scoring to yield reliable results, scorers must adopt the standards of a group of experts. "The face validity of a given test of writing ability depends on whether one agrees with the criteria for judgement established for the ratings." (Charney, 1984, p.73) The validity of writing tests are therefore subject to dispute as conceptualisations of what constitutes good writing may differ from expert to expert.

As the research findings presented earlier illustrate, it is not always possible for scorers to adhere (or agree) to the evaluative criteria specified in holistic grading guidelines. In view of the fact that many teachers of English are inexperienced and untrained in holistic evaluative procedures, we can only assume that the face validity of the evaluative criteria for writing in Unit Curriculum English remains an open question. Reliability thus emerges as a serious problem that could undermine the Ministry of Education's claims of statewide standards and comparability in Unit Curriculum English.

To close this section of the review it is necessary to consider the implications of the research findings presented. First and foremost, there is a clear

indication that holistic scores are significantly affected by the surface usage characteristics of writing. This is interesting considering the fact that these characteristics can be reliably assessed using indirect methods which proponents of the holistic method criticise for being invalid. This fact needs to be carefully considered by testing agencies and education authorities. As Charney points out, the fact that holistic and indirect scores correlate (as supported by Stiggins, 1982; and Veal & Hudson, 1983) "does not establish that neither is valid, but merely that the two tests measure some of the same things... it might mean quantitative measures are more valid than they ought to be or that holistic ratings should be called into doubt." (p.76)

It may prove prudent to digress momentarily to comment briefly on the external procedures used by the SEA in its marking of TEE English and English Literature papers. Notwithstanding the validity criticism put forward by Charney (1984), it is felt that the SEA's procedures for evaluating the TEE English and English Literature papers represent an innovative variation on general impression marking, and an admirable commitment to efficient evaluation. As effective as the procedures are, the situation is atypical, and avoids the problems faced at

school level where there are many inexperienced teachers, inadequately trained in holistic marking procedures who are faced with substantial marking loads. Nevertheless, the SEA's procedures serve as an example that, given certain, and in this case, atypical, conditions, holistic evaluation procedures can be put to good effect.

Discussion

In view of the findings expressed in this paper, it is felt that education authorities are faced with two problems. The first centres around the fact that the existing procedures for the assessment and grading of writing in Unit Curriculum English suffer problems of reliability and validity. These problems clearly extend to include the internal evaluation procedures for the upper school English, English Literature and Senior English subjects.

Synthesis of the research findings presented shows that the factors influencing holistic essay scores challenge the assumption that the direct method of determining writing proficiency is the most valid. In practical terms, we can only assume that problems of this nature are exacerbated at school level, where the teachers of English are invariably untrained in holistic evaluation

procedures and who are faced with syllabus directives which subsume substantial marking loads. This introduces questions of internal reliability as well as the statewide reliability of grades, upon which the Unit Curriculum's standards-referenced evaluation procedures rely. This problem opens questions regarding the credibility of the Ministry of Education's claims on statewide standards in English. Moderation visits can only achieve so much in view of the quantity of work assessed and graded in schools. This problem is real.

Recent linguistic theory applied to the area of writing evaluation has identified the conceptual deficiencies of existing holistic measures (Gere 1980). It is not the purpose of this paper to identify these, as issues of this type steer more toward linguistics and semantics than they do teaching. Nevertheless, the work of Halliday (1978) and Hirsch (1977), seems to indicate that linguistic research provides the most promising avenue to arrive at a concept of meaning in language that can be used as a base for evaluating meaning in writing.

The second problem facing education authorities, then, also applies to teachers. In the absence of a linguistic theory from which a model of evaluating meaning in writing can be developed, current procedures need to be improved. We have no option at this stage but to make full use of what is available. This issue will be discussed in Sections Three and Four of this paper.

Recommendations

As stated in this paper's introduction, this section will close with broad suggestions for further research. As described earlier in more detail, we have yet to determine the full extent of the ramifications of Unit Curriculum structures on English teachers' marking loads. In light of the research findings presented so far, and in the interests of improving the teaching of English, it is recommended that research be funded and initiated to:

1. report, via case-study methodology, the effect of Unit Curriculum time-frames, assessment structures and procedures on the marking loads of teachers of Unit Curriculum English.

It was also stated earlier that indirect methods testing could serve some use in secondary school English. The indirect testing method may prove useful in determining the entrance points of primary school students about to commence Unit Curriculum English. A calculated score that represents some percentage of the indirect test result, together with a score from the existing determination procedures, would yield a representative performance indicator that could be used to decide more

accurately where students enter the English unit stage continuum at secondary school. This possible use is offered in light of the functional skill emphasis of Unit Curriculum English. The analysis of data obtained from these tests would also serve to assist syllabus writers to improve the transition from primary to secondary school English. It is recommended that research be initiated to:

2. investigate the benefits to secondary school decision-makers of data obtained from indirect testing for purposes of determining entrance points of incoming students to Unit Curriculum English.

These recommendations for research will be accompanied by further recommendations in the next two sections. Both sections arrive at conclusions based upon document analysis of the Unit Curriculum syllabus documents synthesised with the research findings presented in this section.

SECTION THREE

A Conceptualisation of Writing Proficiency

Some teachers and parents talk as though there is an agreed global concept of "good writing" (like "intelligence") which everybody can recognize. It's only too clear, though, that notions of good writing (and the criteria for recognizing it) vary from group to group, and for individuals within those groups.

Robert Protherough

In the closing stages of the previous section it was pointed out that conceptualisations of writing proficiency differ from expert to expert. The research presented supported this claim and illustrated that for essay scores to yield statistically reliable results, scorers needed to be well-trained and monitored while marking. This ensured compliance to the established evaluative criteria. The conclusion was made that due to the fact that teachers of Unit Curriculum English are faced with substantial marking loads, and that many are inexperienced and untrained in holistic procedures, the validity of the evaluative criteria for assessing and grading writing in Unit Curriculum English was open to question. Concomitantly, the issue of reliability emerged as a serious problem. Teachers' adherence to the evaluative criteria provided for the assessment and grading of writing in Unit Curriculum English, underpins the Ministry of Education's claim on comparability and statewide standards. In the light of the research and theory presented, the credibility of this claim was questioned.

A conceptualisation of writing proficiency in Unit Curriculum English should reflect syllabus directives and objectives. It was pointed out earlier in this paper that Unit Curriculum English was developed in the light

of recommendations of the Beazley Report (1984). These directives included the call for more emphasis on the development of literacy skill competencies. A conceptualisation of writing proficiency in Unit Curriculum English will acknowledge this directive.

Common to all units of Unit Curriculum English are the process objectives. These objectives describe the range of skills expected to be developed in students, and are central to the evaluative procedures of Unit Curriculum English. Most of these objectives have clear applicability to writing. A conceptualisation of writing proficiency in Unit Curriculum English will acknowledge these objectives also. For ease of reference, the process objectives are re-presented below. The stressed phrases indicate relevance to writing.

- * **use the conventions of standard English in writing;**
- * **prepare and participate in a range of one-to-one, individual and group oral language activities;**
- * **understand, order and convey facts, ideas and opinions in a variety of comprehending and composing situations;**

- * understand and respond to structure, style and tone and vary language according to audience and purpose; and
- * understand and use a wide vocabulary.

These explicit process objectives have been developed in response to the Beazley Report's literacy directive.

There are two additional sources which could be seen to contribute to a conceptualisation of writing proficiency in Unit Curriculum English. Both of these will be rejected. The first of the additional sources are the text objectives. As described earlier, text objectives are task specific, and are supported by the suggested activities presented in the descriptions of each unit. The text objectives and suggested activities reinforce the process objectives of Unit Curriculum English. As such, they play a support role in the determination of what characterises a conceptualisation of writing proficiency in Unit Curriculum English.

The second of the additional sources are the grade-related descriptors (GRDs). GRDs are brief statements designed to supplement the essay scales referred to during the assessment and grading of writing. They form part of the standards-referenced evaluation system.

A Composes writing which:	B Composes writing which:	F Composes writing which:
shows clear expression and direction	is generally clear in expression and direction	displays little sense of direction audience and purpose
regularly obeys most conventions of spelling, punctuation and paragraphing	obeys most conventions of English usage with only minor errors	shows random understandings of the conventions of English usage
shows a variety of vocabulary and sentence types appropriate for the task, and intended audience	reflects clear attempts to vary language for different tasks and audience	carries a negative impact on the reader
is imaginative	is interesting	
carries a strong impact for the reader	has some positive impact on the reader	

Table 4: Grade-related Descriptors for Grades A, B and F for Writing in General Category Units 3.1 and 3.2, and Focus Category Units 3.3 and 3.4

It is clear that the majority of these GRDs reflect the process objectives. Those GRDs which cannot be directly applied, such as; "is imaginative", "carries a strong impact for the reader", and "is interesting", are affective, unqualified and subjective statements. As such, they heighten the problems associated with the validity of the evaluative criteria, and threaten the reliability of grades awarded under Unit Curriculum English. For these reasons, the GRDs are rejected.

A conceptualisation of writing proficiency in Unit Curriculum English, then, will need to be based on the relevant components of the process objectives.

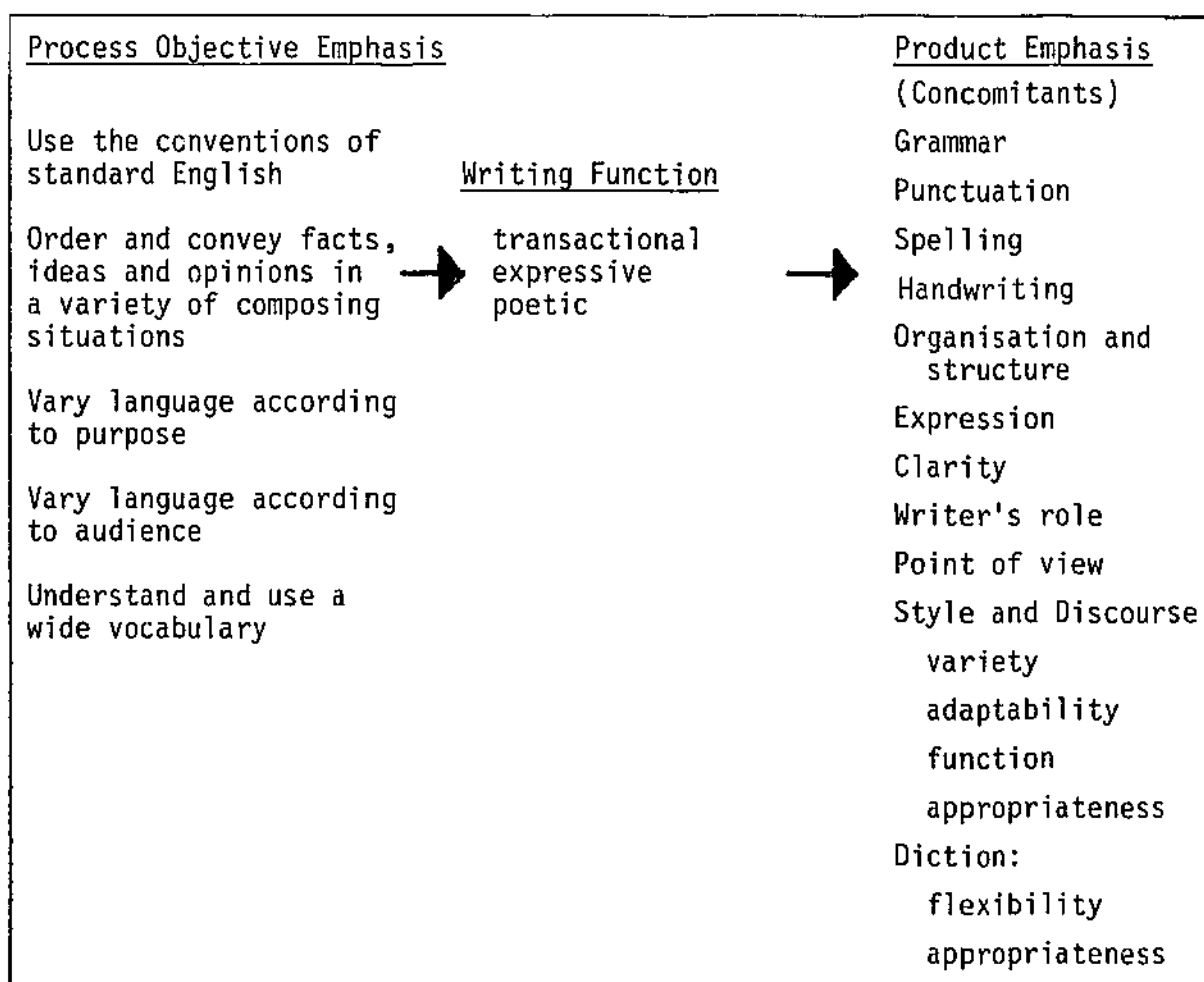


Table 5: A Conceptualisation of Writing Proficiency in Unit Curriculum English

This diagrammatic representation highlights the concomitant skills of the Unit Curriculum English process objectives. The concomitants are, by necessity, highly generalised. They serve to illustrate the wide range and scope of skills that characterise the Unit Curriculum English writing component. The concomitants clearly illustrate that evaluating writing is a complex task.

The evaluative criteria (concomitants), shown in Table 5, also illustrate that there are many factors which need to be considered when assessing and grading writing. The question to then ask is whether the essay scales and grade-related descriptors for Unit Curriculum English adequately address the criteria.

Very little research has focused on the effectiveness of essay scales. Cooper (1977) points out that more efficient and reliable holistic measures have been developed since the advent and widespread use of essay scales in the 1920s and 1930s (p.7). The search of literature conducted for this paper confirmed this view. As the research in this area concentrates on identifying variables which affect the reliability and validity of holistic essay scores, practicality dictates the use of explicit and efficient evaluative methods. Analytic scales, general impression marking and primary trait scoring attract researchers for this reason.

It is possible, nevertheless, to identify certain needs which an evaluative method for writing in Unit Curriculum English should meet. These needs are based on the research findings and theory presented in the preceding section. The extent to which these needs are met will, in part, determine the adequacy of the essay scales and grade-related descriptors used in Unit Curriculum English.

The teachers of Unit Curriculum English need to be provided with an explicit list of evaluative criteria in order to make sound judgements concerning the value of a piece of writing. These explicit criteria would assist in achieving reliability of grades awarded under Unit Curriculum English.

It should be pointed out that essay scales are not explicit statements, but rather a series of complete essays ranging from "A" to "F", accompanied by brief impressionistic remarks. The essays are used as guides. Teachers attempt to match the essays of their students to those in the essay scale and award a grade accordingly. As described earlier, the GRDs which accompany the essay scales in Unit Curriculum English are affective, unqualified and subjective statements. They lack explicitness and threaten the reliability of grades awarded under Unit Curriculum English.

Teachers of Unit Curriculum English need a practical evaluative method. One which is efficient and can serve purposes of formative and summative evaluation. The substantial marking loads confronting teachers of Unit Curriculum English dictate this need.

Essay scales, due to their volume and lack of explicit criteria, require close reading and familiarisation. In this sense, they are impractical. It is often difficult to discern the distinction between an "A" level paper and a "B" level paper in the essay scales for Unit Curriculum English. The remarks accompanying each essay fail to make consistent comparative links between the different levels of response. This vagueness jeopardises the reliability of scores. These criticisms also highlight the point that the formative role of essay scales is limited.

Teachers of English need an holistic evaluation method that requires less training (via virtue of explicitness) and allows for ease of familiarity with its evaluative criteria. This is a practical necessity.

As previously stated, in order that essay scales serve as productive cues to reliable scoring, close reading and familiarisation is required. This implies thorough

discussion of each essay within the scale, and considerable expertise on the markers' behalf to apply these standards to the quantities of work submitted for evaluation. As pointed out by the research, this requires extensive training and monitoring. The realities of teaching practice suggest that these procedures are impractical and idealistic.

Unit Curriculum English requires students to compose writing that covers a wide range of purposes for different audiences. Teachers of Unit Curriculum English therefore need an evaluative method that can be easily adapted to suit all modes of writing across all ability levels.

For essay scales to be truly effective, the scorer needs to be provided with a set of scales which cover the complete range of writing modes across all ability levels. This is not the case in Unit Curriculum English. Essay scales are provided for most General category units. There are no essay scales for Focus or Special Emphasis units. Teachers are directed to refer to essay scales of "similar" units. For example, the essay scales for General category unit 6.1, apply also to General category unit 6.2, and Special Emphasis units 6.3 and 6.4 (Literature), 6.5 (Media), and 6.6 (Innovative Writing). The essay scales for these units

cover only two writing modes: autobiography and short story. This is clearly inadequate.

In the current climate, teachers of English need an evaluative method that, in the eyes of students, is effective, fair and helpful. Teachers should always be in a position to justify their evaluative decisions.

It is difficult to say whether students of Unit Curriculum English find the essay scales and GRDs helpful, fair or effective. Nevertheless, it is the students' right to know how they are being evaluated, and by what criteria their work is being judged. The current procedures pose problems in this respect for the reasons outlined above.

There are two points which bear relevance to the Ministry of Education. The reliability of grades awarded under Unit Curriculum English is central to its claims on statewide standards and comparability. If the evaluative method is unreliable, the Ministry's credibility is at risk.

The Ministry of Education also needs an evaluative method for writing in Unit Curriculum English that is practical and beneficial to its teachers. If an evaluative method is seen by teachers as a compromise, its value will be questioned.

It is not the intent of this paper to deride the assessment and grading procedures for writing in Unit Curriculum English. Nor is it to cast doubt on the integrity on the many professionally minded teachers of English in secondary schools. The evaluation of writing is a clouded sphere within the realm of teaching - it always has been - and the research and theory proves this. Water-tight solutions to the problems of validity and reliability in evaluating writing simply do not exist. The purpose of this paper is to identify these problems and describe how they apply to Unit Curriculum English. The paper also aims to suggest how, if at all, the current assessment and grading procedures for writing can be improved in the light of significant research and related theory.

Before closing this section of the paper, discussion will be made of an evaluative method which offers attractive benefits to the Ministry of Education and teachers of Unit Curriculum English. The evaluative method discussed answers the needs which have been identified.

Analytic Scales

Analytic scales were described in Section Two of this paper. To summarise briefly, an analytic scale is comprised of a list of features common to a writing

mode. Each feature of the scale is categorised under a general trait and divided into three categories: low, middle and high, to which numerical values are attached. The numerical values of each feature vary according to their importance in the mode of discourse being evaluated. Over the years, analytic scales have become more detailed and sophisticated in design. The research of Veal and Hudson (1983) and Wesdorp et al. (1982) indicates that analytic scales are the most reliable of holistic evaluative procedures. Two examples of recently developed analytic scales appear in the Appendix of this paper.

Analytic scales are explicit in that they are comprised of the specific evaluative criteria appropriate to the type of discourse being assessed and graded. In this sense they are useful for achieving reliability of grades. (Veal and Hudson, 1983; Wesdorp et al., 1982.) Liner and Kirby (1981) also point out that as the evaluative criteria of analytic scales are numerically weighted, the surface features of writing will not influence the score out of proportion to the piece's effectiveness. In this sense, analytic scales are a valid means to evaluate writing proficiency.

Analytic scales also serve formative and summative evaluative ends, and can be useful aids in the teaching of writing. Students are provided with a list of points that highlight their writing's strengths and weaknesses (Cooper, 1977).

An analytic scale designed for a particular mode of writing can be validly re-used to assess and grade the same mode of writing of students from all ability levels (Cooper, 1977). In this sense, analytic scales are a practical and economic evaluative method.

As only one analytic scale for each mode of writing prescribed in syllabus objectives is required, the familiarisation of each scale's criteria becomes a practical reality (Cooper, 1977). This lessens the emphasis on training needed to achieve reliable results. This advantage is something not always possible with other, less explicit, evaluative types used to assess and grade writing.

Similarly, the use of analytic scales allows for the reporting of students' abilities on a wide range of traits considered to be important to "good" writing (Stiggins, 1982). In the current climate, this is a major point for consideration.

As analytic scales are explicit in nature and allow for ease of criteria familiarity, teachers can cross-mark papers for internal consistency. This is obviously the first step toward achieving comparability.

Most important, however, is the fact that scorers using analytic scales can be confident of their evaluative decisions. This would do much to reduce the confusion and anxiety associated with assessing and grading writing. The removal of these stresses from the evaluation process would be welcomed by education authorities, and lauded by teachers of English.

Recommendations

In view of the findings expressed in this section of the paper, and in the interests of improving teaching in Unit Curriculum English, it is recommended that:

3. the Ministry of Education develop analytic scales for evaluating writing across all modes of discourse specified in Unit Curriculum English;
4. that these scales be distributed to all secondary schools for purposes of evaluating writing in Unit Curriculum English and be accompanied by detailed notes as to their effective use and benefits to students, teachers, parents and the Ministry of Education;

5. that upon distribution and dissemination of these materials the essay scales and grade-related descriptors be removed from the Unit Curriculum English assessment and grading procedures for writing.

Failing the implementation of these recommendations, the current procedures for assessing and grading writing in Unit Curriculum English need to be improved. These improvements will be discussed in the next section of this paper.

SECTION FOUR

Improving Assessment and Grading Procedures

The previous section of this paper offered a conceptualisation of writing proficiency in Unit Curriculum English. The conceptualisation highlighted the product emphases, or concomitants, of the Unit Curriculum English process objectives. The concomitants illustrated that the evaluation of writing is a complex task. It was pointed out that an evaluative method for writing in Unit Curriculum English should meet the needs of English teachers and the Ministry of Education. A number of needs were identified. These needs were determined by synthesising the concomitants with the research findings and theory presented in Section Two. The question discussed concerned the degree to which the essay scales and grade-related descriptors for Unit Curriculum English addressed these needs. Problems and shortcomings were revealed, and highlighted the need for a more efficient and reliable evaluative method. Analytic scales were accordingly identified as the most appropriate evaluative method for assessing and grading writing in Unit Curriculum English. A criterion-based evaluative method is the most logical choice for the Unit Curriculum's product-based syllabus and standards-referenced evaluation system. Recommendations were made to have the essay scales and grade-related descriptors replaced by the more efficient and reliable analytic scales. It was pointed out that failing the

implementation of these recommendations, the existing guidelines and procedures for assessing and grading writing in Unit Curriculum English needed to be improved.

This section of the paper is brief, and will identify aspects of the guidelines and procedures for assessing and grading writing in Unit Curriculum English which need clarification and improvement. Recommendations will be made accordingly. The points raised in this section stem largely from the criticisms of the essay scales and grade-related descriptors discussed in detail in Section Three. Underlying the points is the belief that in order to reduce the problems of reliability and validity associated with evaluating writing, teachers of Unit Curriculum English need explicit guidelines and procedures in order to make sound evaluative decisions.

It was pointed out earlier that in order for essay scales to be truly effective, scorers need to be provided with a set of scales which cover the complete range of writing modes across all ability levels. An example was provided to show that this was not the case in Unit Curriculum English. The major point of concern centred around the fact that the essay scales failed to cover adequately the writing modes required to be assessed and graded.

It was also mentioned that the summary statements accompanying the essays within the scales lacked explicitness. They failed to draw consistent comparisons among the essays and, in doing so, failed to distinguish clearly the difference between the levels of graded response.

The grade-related descriptors are designed as supplemental sources to assist teachers in the assessment and grading of writing. It was pointed out that the majority of these were reflections of the explicit process objectives. There were, however, a number of grade-related descriptors not directly attributable to the process objectives. These were criticised for being affective, unqualified and subjective. It was felt that these grade-related descriptors threatened the reliability of scores upon which the Ministry of Education's claims on comparability and statewide standards in Unit Curriculum English rely. In view of these criticisms it is recommended that:

6. the Ministry of Education review the current guidelines and procedures for assessing and grading writing in Unit Curriculum English with due consideration to the recommendations to follow;

7. the essay scales for Unit Curriculum English be made more comprehensive to include all relevant writing modes for each unit of Unit Curriculum English;
8. the summary notes in the essay scales be re-written to identify the strengths and weaknesses of each essay together with clear reasons as to what distinguishes each essay from others in the scale;
9. the grade-related descriptors not directly attributable to the process objectives be re-written to be more explicit and detailed with qualified examples so as to reduce subjectivity and ambiguity.

The standards expected today may not be indicative or relevant to the standards required tomorrow. For this reason, it is recommended that:

10. the essay scales and grade-related descriptors for Unit Curriculum English be regularly assessed for appropriateness so as to reflect the changing standards of writing proficiency expected of students in secondary schools.

One of the problems facing teachers of Unit Curriculum English concerns the bulk of material needed to be referred to while using the essay scales and grade-related descriptors. At present, the documents concerned total 258 pages. The essay scales appear in one document, the grade-related descriptors in another. This is inconvenient. The sheer bulk of documents present a more serious problem, ironic in nature. If essay scales are used to their full extent, as is recommended here, the bulk of reference material would be substantially increased. This is an inherent problem of essay scales. Nevertheless, if the summary statements of each essay of an essay scale were accompanied by a list of GRDs, this problem would be reduced. It is surely more convenient to have one document containing all the necessary criteria, rather than having two documents sharing them. If a system of reference is meant to be used, it should be designed for ease of user reference. It is recommended that:

11. the format, design and packaging of essay scales and grade-related descriptors for Unit Curriculum English be modified to serve reasons of practicality and ease of reference.

There are some additional concerns with the existing guidelines and procedures which have not previously been alluded to. These concerns reflect the lack of explicit requirements pertaining to the quantity of writing students are to submit for evaluation. At present, issues of this nature are left for schools to resolve. It is not unreasonable to state, then, that different schools require different quantities of writing from their students. It is felt that in order to guard against "over" or "under" evaluation, and to achieve statewide standards, the Ministry of Education should specify the quantities of writing required of students in all units of English.

So as to ensure all aspects of writing are adequately covered during the course of a unit, explicit requirements are needed to clarify the number of written pieces students are to submit from each text objective heading. A grade in English should be representative of a student's ability to write proficiently in all modes of writing. Competence in one mode of writing is not a pre-cursor of competence in another. The current guidelines and procedures fail to address this concern explicitly. In the interests of statewide standards and efficient evaluation, it is recommended that:

12. the Ministry of Education specify for each unit of English the number of written pieces students are to submit for purposes of summative evaluation from each text heading of the syllabus.

It has been mentioned that the reliable grading of writing in Unit Curriculum English is central to the Ministry of Education's claims on comparability and statewide standards. It was also mentioned that checks for internal consistency were the first step toward achieving comparability. Many English Departments encourage and practice cross-grading, and meet regularly to discuss standards so as to allow for consistent and reliable grading. Many do not. For this reason, the Ministry of Education should make rigorous attempts to ensure that the criteria for the assessment and grading of writing in Unit Curriculum English be made explicit.

Efforts should also be made to ensure comparability between schools. The evaluative criteria for writing in Unit Curriculum English need to be consistently applied in all schools. This highlights the need for extensive moderation procedures. In the interests of English teaching, comparability and statewide standards, it is recommended that:

13. the Ministry of Education review moderation procedures in Unit Curriculum English and encourage regular comparability checks within and between schools throughout the State.

The problems identified here concerning the guidelines and procedures for assessing and grading writing question the credibility of the Ministry of Education's claims on comparability and statewide standards in Unit Curriculum English. The recommendations offered represent efforts to address the question.

Closing Statement

It was stated earlier that the evaluation of writing is problematic. In a system where evaluation is given particular emphasis, these problems are exacerbated. This heightens the need for extensive and systematic reviews of the guidelines and procedures used. A system should always be under review. Monitoring standards is not enough. The need to address the insufficiencies of the guidelines and procedures for assessing and grading writing in Unit Curriculum English is drastic. Teachers are held accountable for their actions. Their evaluative, and hence, public, decisions, are at present executed in compliance to the guidelines and procedures described in this paper. The paper has identified that these guidelines and procedures are lacking. This can reflect poorly upon the teachers of Unit Curriculum English.

While it may be politically expedient to lay claim to comparability and statewide standards in Unit Curriculum English, the point needs to be stressed that the current guidelines and procedures for assessing and grading writing do not adequately address the requisites of the claim. It is for this reason also the insufficiencies of the guidelines and procedures need to be addressed. Failure to do so could result in teachers of English being placed in a position susceptible to public and political criticism.

List of Recommendations

1. Report, via case-study methodology, the effect of Unit Curriculum time-frames, assessment structures and procedures on the marking loads of teachers of Unit Curriculum English.

2. Investigate the benefits to secondary school decision-makers of data obtained from indirect testing for purposes of determining entrance points of incoming students to Unit Curriculum English.

3. The Ministry of Education develop analytic scales for evaluating writing across all modes of discourse specified in Unit Curriculum English.

4. That these scales be distributed to all secondary schools for purposes of evaluating writing in Unit Curriculum English and be accompanied by detailed notes as to their effective use and benefits to students, teachers, parents and the Ministry of Education.

5. That upon distribution and dissemination of these materials the essay scales and grade-related descriptors be removed from the Unit Curriculum English assessment and grading procedures for writing.

Failing the implementation of recommendations 3-5, the following recommendations are offered.

6. The Ministry of Education review the current guidelines and procedures for assessing and grading writing in Unit Curriculum English with due consideration to recommendations 7, 8 and 9.
7. The essay scales for Unit Curriculum English be made more comprehensive to include all relevant writing modes for each unit of Unit Curriculum English.
8. The summary notes in the essay scales be re-written identifying the strengths and weaknesses of each essay together with clear reasons as to what distinguishes each essay from the others in the scale.
9. The grade-related descriptors not directly attributable to the process objectives be re-written to be more explicit and detailed with qualified examples so as to reduce subjectivity and ambiguity.

10. The essay scales and grade-related descriptors for Unit Curriculum English be regularly assessed for appropriateness so as to reflect the changing standards of writing proficiency expected of students in secondary schools.

11. The format, design and packaging of essay scales and grade-related descriptors for Unit Curriculum English be modified to serve reasons of practicality and ease of reference.

12. The Ministry of Education specify for each unit of English the number of written pieces students are to submit for purposes of summative evaluation from each text heading of the syllabus.

13. The Ministry of Education review moderation procedures in Unit Curriculum English and encourage regular comparability checks within and between schools throughout the State.

Appendices

From:

Cooper, C.R., & Odell, L.
Evaluating Writing. Describing,
Measuring, Judging. Ohio:
NCTE, 1977

APPENDIX A: A Personal Narrative Writing Scales

1. General Qualities:

A. Author's Role

The author's role is the relationship of the author to the subject, incident, or person. In **autobiography** the author writes about himself/herself. He/she is the main participant. Most of the time he/she will use the pronouns, I, me, we us. In **biography** the author writes about some other person. He/she is not involved in what happens; he/she is just an observer. He/she uses the pronouns, he, she, him, her, it, they, them.

High The author keeps his/her correct role of either participant or observer throughout.

Middle In autobiography, a few noticeable distracting times the author talks too much about another person's actions; or, in biography, he/she talks too much about his/her own actions.

Low The author talks about himself/ herself or others as participant or observer anytime he/she pleases so that you can barely tell whether it is supposed to be autobiography or biography. There is confusion as to author's role. He/she is not consistently either observer or participant.

B Style or Voice

High The author states what he/she really thinks and feels. Expressing personal experiences, the writer comes through as an individual, and his/her work seems like his/hers and his/hers alone. The voice we hear in the piece really interests us.

Middle The author uses generalizations or abstract language, seldom including personal details and comments. While the piece may be correct, it lacks the personal touch. The voice seems bland, careful, a little flat, and not very interesting.

Low We don't really hear a recognizable voice in the piece. The style seems flat and lifeless.

C. Central Figure

Details about the central figure make him/her seem "real". The character is described physically and as a person.

High The central figure is described in such detail that he/she is always "real" for you.

Middle The central character can be "seen," but is not as real as he/she could be.

Low The central character is not a real living person; he/she is just a name on a page. You cannot see him/her or understand him/her.

D. Background

The setting of the action is detailed so that it seems to give the events a "real" place in which to happen.

High The action occurs in a well-detailed place that you can almost see.

Middle Sometimes the setting seems vivid and real; but sometimes the action is just happening, and you are not really aware of what the setting is.

Low The action occurs without any detailed setting. You see the action, but you cannot see it in a certain place.

E. Sequence

The order of events is clear, giving the reader a precise view of the sequence of incidents.

High The order of events is always clear to you even if at times the author might talk about the past or the future.

Middle A few times it is not clear which event happened first.

Low You really cannot figure out which event comes first or goes after any other event.

F. Theme

The author chooses the incidents and details for some reason. There seems to be some purpose behind the choice of subject matter, some theme holding it all together and relating the parts to the whole. There seems to be a point to it.

High The importance of the author's subject is either directly explained to you or it is implied in a way that makes it clear.

Middle You can see why the author's subject is important to him/her, but it is not as clearly stated or implied as it could be.

Low You cannot figure out why the subject is important to the author.

II. Diction, Syntax, and Mechanics

A. Wording

High Words are employed in a unique and interesting way. While some of the language might be inappropriate, the author seems thoughtful and imaginative.

Middle Common, ordinary words are used in the same old way. The paper has some trite, over-worked expressions. The author, on the other hand, may work so hard at being different that he/she sounds like a talking dictionary, in which case he/she also, merits this rating.

Low The word choice is limited and immature. Sometimes words are even used incorrectly - the wrong word is used.

B. Syntax

High The sentences are varied in length and structure. The author shows a confident control of sentence structure. The paper reads smoothly from sentence to sentence. There are no run-together sentences or sentence fragments.

Middle The author shows some control of sentence structure and only occasionally writes a sentence which is awkward or puzzling. Almost no run-ons and fragments.

Low Many problems with sentence structure. Sentences are short and simple in structure, somewhat childlike and repetitious in their patterns. There may be run-ons and fragments.

C. Usage

High There are no obvious errors in usage. The author shows he/she is familiar with the standards of edited written English.

Middle A few errors in usage appear in the paper, showing the author has not quite been consistent in using standard forms.

Low The writing is full of usage errors.

D. Punctuation

High The author consistently uses appropriate punctuation.

Middle Most of the time the writer punctuates correctly.

Low The writing contains many punctuation errors.

E. Spelling

High All words are spelled correctly.

Middle A few words are misspelled.

Low Many words are misspelled.

Analytic Scale

	Reader			Paper	
	Low	Middle	High		
I. General Qualities:					
A. Author's Role	2	4	6	8	10
B. Style or Voice	2	4	6	8	10
C. Central Figure	2	4	6	8	10
D. Background	2	4	6	8	10
E. Sequence	2	4	6	8	10
F. Theme	2	4	6	8	10
II. Diction, Syntax, and Mechanics:					
A. Wording	1	2	3	4	5
B. Syntax	1	2	3	4	5
C. Usage	1	2	3	4	5
D. Punctuation	1	2	3	4	5
E. Spelling	1	2	3	4	5
				Total	

APPENDIX B: Dramatic Writing Scales

The language of dramatic writing is different from other types of writing because it is meant to be heard. We expect the language to be in the present tense because the events unfold as we watch and listen. Another special aspect of the language of dramatic writing is that there is no narrator or voice to tell us of descriptions and histories. In dramatic language this information is hidden in the face-to-face, ongoing conversations of the characters. While each character speaks, other tenses than the present are used to talk to other characters. For example, one character may relate to another a past series of events leading to the present situation. The stage directions give hints to the actors concerning their actions and tone of voice, which the narrator would otherwise tell about in a descriptive section of prose.

I. Language Factors

A. Conversation: Realism Does the conversation sound realistic?

High The characters' conversations go on as if you were eavesdropping (secretly listening) to their talk. Everything that is said is very clear to you.

Middle The characters' conversation sometimes leaves out something important. Almost everything that is said is clear to you.

Low The characters' conversation leaves out so much that you have trouble understanding what is said.

B. Conversation: Situation Does the way the characters talk match the situation they are in?

High The characters talk exactly as you would expect in the situation.

Middle The characters talk as you expect in the situation most of the time.

Low The characters do not talk as you would expect in the situation.

C. Stage Directions

If stage directions are used, are they short and clear?

High The stage directions tell the actors how to act and speak when you cannot decide from the characters' talk.

Middle The stage directions tell the actors how to act and speak most of the time. Sometimes they leave information out or repeat information.

Low The stage directions confuse the actors about how to act and speak.

II. Shape Factors

A. Beginning

Does reading the opening lines of this dramatic writing make you want to continue? Do they make you feel that what follows will be interesting?

High I am intrigued by the beginning. It seems interesting and makes me want to continue the reading.

Middle The beginning is interesting; however, I have seen this beginning used before. It's not all that unusual.

Low The beginning is not particularly interesting. It gets the dramatic writing off to a slow start.

B. Structure

Structure refers to the way this dramatic writing is built, or put together, with a beginning, middle, and end. It has to do with the way the parts fit together, the overall design which reveals the problem and how that problem is solved.

High The elements of the dramatic writing are tied together in an interesting, well-organized manner. There is a good deal of detail and a resolution that is believable.

Middle Although there is some attempt at proceeding from beginning to end in an organized manner, you are unsatisfied. This could be due either to a "forced" conclusion to the writing or to the writer's failure to tie all the elements together very successfully.

Low The sequence of events is confused, rambling, not well-organized. Very little detail is given.

C. Ending

The ending is the dramatic writing's conclusion. It is reached **after** a problem has been resolved.

- High** The ending follows sensibly from the story, is unique, very well stated, and, possibly, is a surprise ending.
- Middle** The ending makes sense to the dramatic writing but is not very unique or unusual.
- Low** Very ordinary and usual. The ending is just what you expected and does not surprise you. It may not resolve the problem posed in the writing, or it may not resolve it in a believable manner.

III. Characterization Factors

Having characters that are well-developed and real to the readers is an important part of dramatic writing. Making the reader understand how and why the characters act the way they do will give the reader a more personal and interesting view of the entire dramatic piece.

A. Development and Credibility

All the characters in the writing should be as much like real people as possible. The reader should be able to see the difference between the major and minor characters. Major characters (the important ones) should be more fully developed. The reader should know a lot about them. They should see him/her acting and reacting in many different situations. Minor characters (less important ones) also have to be realistic, but the reader doesn't have to know as much about them.

- High** All major characters seem to be like real people. Each character is a different person, and the reader has no problem telling which character is which. Minor characters are also real, but they aren't as detailed as major ones. The writer tells the reader much about his characters through dialogue. Narration is kept to a minimum.
- Middle** Not all the characters seem like real people, all the time. Sometimes they do things that real people probably wouldn't do. The reader has a hard time telling what characters are which. They all seem alike.
- Low** Little about the characters seems real. They act in ways which most people wouldn't. There is no difference between major and minor characters. The characters are almost entirely described by narration, with little use of dialogue.

B. Consistency
The characters seem like the same people throughout the piece of writing. Their emotions might change (they may change by laughing, crying, feeling happy or sad, etc.) but their basic personalities will remain the same. (A boy who was very stingy with his money at the beginning of the story wouldn't suddenly start giving money away for no reason.)

High All the characters remain the same throughout the piece. Their personalities do not change. If there is a basic change, a reason is given for it in the dialogue.

Middle The characters do not always seem like the same people. There are times when they do things that don't seem to fit.

Low The characters' personalities are constantly changing. The reader never knows what to expect from them.

IV. Mechanics Factors

A. Dramatic Form

Dramatic Form refers to the physical arrangement of words on the paper. Is the physical form of the paper such that the reader wants to continue reading? The names of the characters should come before their lines, and be set off to the left, followed by a colon. If stage directions are used, they should be enclosed in parentheses.

High The form is nearly perfect; stage directions are set off by parentheses.

Middle There are a few errors in form or occasionally confusing stage directions.

Low The paper contains many errors in dramatic form: characters' names are omitted or put in the wrong places. Stage directions are run into the characters' lines.

B. Spelling

Dialect spellings are permitted in dramatic writing. Where they are used, they should be consistent so that the actor would have no difficulty reading the character's lines.

- High** All words are spelled correctly, even the most difficult words. Dialect spellings are consistent.
- Middle** Only a few words are misspelled. Dialect spellings are mostly consistent.
- Low** There are many misspellings, even of very ordinary words. Dialect spellings are inconsistent.

C. End Punctuation

- High** End punctuation occurs at natural places, thus making the dialogue easy to follow.
- Middle** There are only a few errors in end punctuation, without making the dialogue difficult to follow.
- Low** End punctuation marks are either not present or are placed so that often the dialogue is hard to follow.

V. Response Factors

Rather than focusing your attention on one aspect of drama, in this section of the scale you will be asked to assess the dramatic work as a whole. The questions under this heading of the scale will probably be the easiest for you to answer because you know what you like and dislike. However, you should try to use your answers to these questions as guides in answering the other more specific questions. For example, if you really enjoyed a work, try to decide what aspect of the work made it so successful.

A. Entertainment

- High** I felt the work was very entertaining.
- Middle** I was only mildly entertained by the work as a whole.
- Low** The work was not entertaining.

B. Originality

- High** The work made me think about something in a way that I hadn't previously considered.
- Middle** While there were some moments of originality in the work, there were a lot of ideas I had heard before.
- Low** There was nothing new in this work.

Reader Score Sheet

DIRECTIONS: For each quality listed below, circle the number that most nearly describes the position of this paper on the following scale from high to low.

LANGUAGES FACTORS		HIGH	MIDDLE	LOW		
I.1	Conversation - Realism...	5	4	3	2	1
I.2	Conversation - Situation.	5	4	3	2	1
I.3	Stage Directions.....	5	4	3	2	1

TOTAL LANGUAGE SCORE _____

SHAPE FACTORS

II.1	Beginning.....	5	4	3	2	1
II.2	Structure.....	5	4	3	2	1
II.3	Ending.....	5	4	3	2	1

TOTAL LANGUAGE SCORE _____

CHARACTERIZATION FACTORS

III.1	Development	5	4	3	2	1
III.2	Consistency	5	4	3	2	1

TOTAL CHARACTERIZATION SCORE _____

MECHANICS FACTORS

IV.2	Dramatic Form	5	4	3	2	1
IV.2	Spelling	5	4	3	2	1
IV.3	Punctuation	5	4	3	2	1

TOTAL MECHANICS SCORE: _____

RESPONSE FACTORS

V.1	Entertainment	5	4	3	2	1
V.2	Originality	5	4	3	2	1

TOTAL RESPONSE SCORE: _____

TOTAL SCORE : _____

BIBLIOGRAPHY

- Bamberg, B. Assessing Coherence: A reanalysis of Essays Written for the National Assessment of Education Progress, 1969-1979. Research in the Teaching of English, 1984, 18(3), 305-319.
- Beazley, K. (Chairman). The Report of the Committee of Inquiry Into Education in Western Australia. Education in Western Australia. Perth, Western Australia, 1984.
- Biggs, J., & Telfer, R. The Process of Learning. Australia: Prentice Hall, 1981.
- Breland, H.M., & Jones, R.J. Perceptions of Writing Skills. Written Communication, 1984, 1(1), 101-119.
- Britton, J.N., Martin, N., & Rosen, H. Multiple Marking of Compositions. London: HMSO, 1966.
- Butler, S.J. New Bottles for New Wine. Evaluation in a Modern Writing Program. English Quarterly, 1985, 18(2), 57-68.
- Carlman, N. Topic Differences on Writing Tests: How Much Do They Matter? English Quarterly, 1986, 19(1), 39-49.
- Charney, D. The Validity of Using Holistic Scoring to Evaluate Writing: A Critical Overview. Research in the Teaching of English, 1984, 18(1), 65-81.
- Cooper, C.R., & Odell, L. Evaluating Writing: Describing, Measuring, Judging. USA: NCTE, 1977.

BIBLIOGRAPHY

- Culpepper, M.M., & Ramsdell, R.A. Comparison of a Multiple Choice and an Essay Test of Writing Skills. Research in the Teaching of English, 1982, 16(3), 295-297.
- Daly, J.A., & Dickson-Markman, F. Contrast Effects in Evaluating Essays. Journal of Education Measurement, 1982, 19(4), 309-315.
- De Shields, S.M., Hsieh, H.K., & Frost, D. The Measurement of Writing Skills: Some Problems and a Probable Solution. Educational and Psychological Measurement, 1984, 44, 101-112.
- Eagleson, R.D. (Ed.) English in the Eighties. Australia: AATE, 1982.
- Freedman, S.W. How Characteristics of Student Essays Influence Teachers' Evaluations. Journal of Educational Psychology, 1979, 71(3), 328-338.
- Freedman, S.W. Influences on Evaluators of Expository Essays: Beyond the Text. Research in the Teaching of English, 1981, 15(3), 245-255.
- Freedman, S.W. Some Reasons for the Grades We Give Compositions. English Journal, 1982, 71(7), 86-89.
- Freedman, S.W. Student Characteristics and Essay Test Writing Performance. Research in the Teaching of English, 1983, 17(4), 313-325.

BIBLIOGRAPHY

- Gay, L.R. Educational Measurement and Evaluation. Ohio: Charles Merrill Publishing Company, 1985.
- Gere, A.R. Written Composition: Toward a Theory of Evaluation. College English, 1980, 42(1), 44-58.
- Grobe, C. Syntactic Maturity, Mechanics, and Vocabulary as Predictors of Quality Ratings. Research in the Teaching of English, 1981, 15(1), 75-85.
- Hake, R., & Williams, J.M. Style and its Consequences: Do as I Do, Not as I Say. College English, 1981, 43(5), 433-451.
- Halliday, M.A.K. Language as Social Semiotic. Baltimore: University Park Press, 1978.
- Harris, W. Teacher Response to Student Writing: A Study of the Response Patterns of High School English Teachers to Determine the Basis for Teacher Judgment of Student Writing. Research in the Teaching of English, 1977, 11, 175-185.
- Hay, J. General Impression Marking: some Caveats. English in Australia, (59), 50-58.
- Hirsch, E.D. The Philosophy of Composition. Chicago: University of Chicago Press, 1977.
- Hughes, D.C., Keeling, B., & Tuck, B.F. The Effects of Instructions to Scorers Intended to Reduce Context Effects In Essay Scoring. Educational and Psychological Measurement, 1983a, 43, 1047-1050.

BIBLIOGRAPHY

- Hughes, D.C., Keeling, B., & Tuck, B.F. Effects of Achievement Expectations and Handwriting Quality on Scoring Essays. Journal of Educational Measurement, 1983b, 20(1), 65-70.
- Hughes, D.C., & Keeling, B. The Use of Model Essays to Reduce Context Effects in Essay Scores. Journal of Educational Measurement, 1984, 21(3), 277-281.
- Johnston, K. A Discourse for All Seasons? An Ideological Analysis of the Schools Commission Reports, 1973 to 1981. Australian Journal of Education, 1983, 27(1), 17-32.
- Liner, D., & Kirby, T. Inside Out. Developmental Strategies for Teaching Writing. USA: Boynton Cook, 1981.
- Marginson, S. The Collapse of the 1973 Karmel Consensus. AFT Research Papers, 1985, (9).
- Markham, L.R. Influences of Handwriting Quality on Teacher Evaluation of Written Work. American Educational Research Journal, 1976, 13, 277-288.
- Martin, N., Britton, J.N., & Rosen, H. Assessing Compositions. Glasgow: Blackie, 1965.
- McColly, W. What Does Educational Research Say About the Judging of Writing Ability? The Journal of Educational Research, 1970, 64(4), 149-156.

BIBLIOGRAPHY

- McKinnon, K.R. The Schools Commission: Policies and Politics in a Statutory Body. In G. Harman & D. Smart (Eds.), Federal Intervention in Australian Education. Melbourne: Georgian House, 1982.
- Nail, P., Fitch, R., Halverson, J., Grant, P., & Winn, N.F. The Scale for Evaluation of High School Student Essays. USA: NCTEP, 1960.
- Odell, L., & Cooper, C.R. Procedures for Evaluating Writing: Assumptions and Needed Research. College English, 1980, 41(1), 35-43.
- Popham, W.J. Modern Educational Measurement. New Jersey: Englewood Cliffs, 1981.
- Protherough, R. Encouraging Writing. London: Methuen, 1983.
- Stewart, M.F., & Grobe, C. Syntactic Maturity, Mechanics of Writing and Teachers' Quality Ratings. Research in the Teaching of English, 1979, 13, 207-215.
- Stiggins, R.J. An Analysis of Direct and Indirect Writing Assessment Procedures. Evaluation in Education, 1982, 5, 347-357.
- Sweedler-Brown, C.O. The Influence of Training and Experience on Holistic Essay Evaluations. English Journal, 1985, 74(5), 49-55.
- Taylor, S. Ideology, Policy and the Education of Girls. Unpublished monograph, BCAE, 1982.

BIBLIOGRAPHY

- Tierney, R.J., & Mosenthal, J.H. Cohesion and Textual Coherence. Research in the Teaching of English, 1983, 17(3), 215-229.
- Veal, L.R., & Hudson, S.A. Direct and Indirect Measures for Large Scale Evaluation of Writing. Research in the Teaching of English, 1983, 17(3), 290-296.
- Verhulst, D.J. "When to my Eyes I Saw the One I Loved": Applying the Wilkinson Scales to Student Writing. Canadian Journal of English and Language Arts, 1987, 11(3), 31-37.
- Walshe, R.D., March, P., & Jensen, D. Writing and Learning in Australia. Melbourne: Dellasta & OUP, 1986.
- Wesdorp, H., Bauer, B.A., & Purves, A.C. Toward a Conceptualisation of the Scoring of Written Composition. Evaluation in Education, 1982, 5, 299-315.
- Wilkinson, A. Assessing Language Development. England: OUP, 1980.
- Wirt, F. Professional Versus Lay Conflict. A Paradigm for Policy Making in Modern Systems. Edited transcript of seminar, Department of Education. Perth, 1979.