



This is the post-print of Shay, S. 1997. Portfolio assessment: a catalyst for staff and curricular reform. *Assessing Writing*. 4(1): 29-51.

It is made available according to the terms of agreement between the author and the journal, and in accordance with UCT's open access policy available:

<http://www.openuct.uct.ac.za/sites/default/files/UCTOpenAccessPolicy.pdf>, for the purposes of research, teaching and private study.

Portfolio Assessment: A Catalyst for Staff and Curricular Reform

SUELLEN SHAY

University of Cape Town

South Africa's gradual transition from apartheid to democracy has resulted in a steady increase in black enrolment over the year-black students will soon be a majority at the University of Cape Town (UCT). It is a well-known fact that most of these students due to the legacy of apartheid education enter the university highly underprepared and that significant improvements in the educational system are not anticipated in the near future. This demographic transition has forced Academic Development (AD) at UCT to shift its role from supporting a minority of students to developing programmes which cater to the needs of all of UCT's population-both students and staff. The Writing 'Centre's evolving *modus operandi* centers around the assessment of students' writing, in the context of academic departments, as a tool for evaluating existing curricula and pedagogical approaches. This article explores how a portfolio assessment project in the Chemistry Department developed (within staff) new understandings of the role of writing as a vehicle for learning, which resulted in a re-examination of the relationship of the written assignments to the existing curriculum. Based on our experience, I offer in the conclusion some suggestions of ways in which the potential of portfolio assessment for programme evaluation may be exchanged given further research. Specifically, the paper discusses the challenges of designing assessment frameworks which are diagnostically powerful enough to assist us in the task of curriculum reconstruction.

THE NATIONAL AND INSTITUTIONAL CONTEXT: A TRANSFORMATIVE AGENDA

David Russell (1991), in his analysis of the American writing-across-the-curriculum movement, states that "cross-curricular writing programs were almost always a response to a perceived need for greater access, greater equity. They (these programs) set out to assimilate, integrate, or initiate previously excluded students by means of language instruction" (p. 271). To some extent this aptly describes the *raison d'être* of Academic Development (AD) work when it was initiated in the early 1980s in South Africa's historically white universities. The aim was to "assimilate" a small number of black' students who, due to the legacy of apartheid education, were underprepared for the demands of tertiary education. Not only had their schooling failed to provide them with the necessary academic foundations, but they confronted the reality of English as the exclusive medium of instruction. In addition, many of them experienced tremendous cultural disorientation in what was then an overwhelmingly white, western, middle-class, urban environment.

South Africa's gradual transition from apartheid to democracy has resulted in a steady increase in black enrolment over the years-black students will soon be a majority at the University of Cape Town (UCT). The effects of apartheid education are as devastating in 1995 as they were a decade ago and significant improvements in the educational system are not anticipated in the near future. This demographic transition has forced AD at UCT (and at other historically white universities) to shift its role from "*supporting a minority of black students to systematic and system-wide provision that cater for all of UCT's student population*" (Moore, 1996, p.1; Yeld, 1992). This implies the necessity for departments, with assistance from language and learning experts, to take greater responsibility for providing a learning environment that caters for the needs of a very diverse student population.

In this context the Writing Centre at UCT, while providing the traditional student consultancy service, sees its *ultimate* aim to contribute to institutional transformation by assisting in the creation and implementation of policy which values and rewards curriculum and staff development. The Writing Centre's evolving *modus operandi* centres around close collaboration with academics in the assessment of their students' writing as a catalyst for the development of learning and teaching environments which are more sensitive to the needs of a diverse student population.

Unlike the American context, there are currently no formal writing course requirements necessary for graduation from a South African university. English departments in South Africa have not played the same kind of service role as their American counterparts by providing composition courses and writing-across-the-curriculum programmes. The environment in which we work resembles that described by Lovitt and

Young (1994) where the teaching of writing in the disciplines has to happen in close collaboration with disciplinary academics who are unlikely to share the assumptions and interests of the writing-across-the-curriculum movement. Although this is challenging, we believe that these collaborations are fertile soil for fresh insights on the role of writing in the curriculum and for innovative forms of educational practice.

The aim of this article is to illustrate through writing centre practice the power of portfolio assessment as a catalyst for reform. Specifically, this paper describes how a portfolio assessment project in a chemistry department developed (within staff) new understandings of the role of writing as a vehicle for learning, which resulted in a re-examination of the relationship of the written assignments to the existing curriculum. Based on our experience I offer in the conclusion some suggestions of ways in which the potential of portfolio assessment for programme evaluation may be enhanced given further research.

THE DEPARTMENTAL CONTEXT: WRITING IN CHEMISTRY

In response to a concern from employers that UCT Science graduates were technically competent but unable to communicate this competence in the workplace (Allie, Kaunda, & Holmes, 1995), in 1995 UCT's Chemistry Department initiated a Writing Project—a collaboration between writing specialists from the Writing Centre, consisting of myself and a colleague with a Chemistry background, and discipline specialists from the Chemistry Department. The goal of the Writing Project was, within the context of the second year Chemistry curriculum, to give students genuine opportunities to consolidate, develop and communicate their scientific understanding by assigning three short scientific reports over the course of the year.

Certain assumptions informed the design and mediation of each of these reports, henceforth referred to as the Caffeine, Chlorine and Oxalyl Chloride reports. First, it was important that the genre of the writing assignments be appropriate to the discipline; that is, that students be exposed to writing tasks that bear some resemblance to the communicative tasks in which they will engage as professionals. Related to genre, students need to experience writing for multiple audiences since the communicative skills necessary to write for a fellow-scientist differ from those necessary for an educated person with no scientific training. Second, since for many students these reports would be their first experience of extended writing at a university, the writing assignments should be “scaffolded,” with intensive support on the first report, moving towards greater independence by the third report. (This did not always happen in practice). Third, I assumed that the conceptual content of the reports would be an extension of material covered in the lectures. (I discovered that this latter assumption was not necessarily shared by my colleagues in the Chemistry Department.)

In the first report—the Caffeine report—students were given raw data from an imaginary analytical determination of caffeine in various beverages. Students were to write a report following the traditional format of the lab report. The science necessary for this report was fairly straightforward. Students were required to (1) plot the given data on a calibration curve, and (2) determine the concentration of caffeine in four sample beverages (involving a simple mathematical procedure). In the discussion, students were required to evaluate a number of methods used to decaffeinate coffee beans, and identify why one particular method was preferred by coffee producers. In a two-hour workshop, students were taken through the demands of the task and given detailed input on the lab report format.

The second report—the Chlorine report—was of a very different nature. Students were to play the role of scientific adviser to South Africa's Minister of Science and Technology (informed, educated, but not a chemist). They were required to (1) discuss how chlorine was produced, (2) outline ways in which the chlorine industry contributes to the economy, and (3) discuss some of the environmental considerations of the expansion of such an industry. Students should have been familiar with the production and properties of chlorine from first year Chemistry but had not been exposed to the industrial processes or the environmental issues concerned. Students were supplied with one reading from a popular science journal on the competing interests of industry, government and environment on the production of chlorine in the United States. Although students were not required to take a definitive position, there was an argumentative element in which students needed to consider the various positions of competing stakeholders. Mainly for reasons of timetabling constraints, the students were given no assistance on this report.

The third report—the Oxalyl Chloride report—was also of a different nature. Students were to write a report to a company Managing Director (with a Ph.D. in synthetic organic Chemistry) on the advantages of a

recently developed modified procedure for the production of oxalyl chloride from ethene carbonate. As in the Caffeine report, they had to evaluate the advantages of this process against other methods. As in the Chlorine report, they were asked to consider the environmental aspects. But unlike the previous reports, the Oxalyl Chloride report required knowledge and use of chemical language. This was by far the most conceptually challenging report. The analytical principles underlying the task had been presented to students in the previous semester but transfer of this information to the assignment was not explicitly made. In a half-hour presentation by the lecturer, students were taken through the demands of this task.

THE PROBLEM: HOW TO ASSESS CURRICULAR?

It might have been tempting to conclude that as a result of the Writing Project second year Chemistry students were better prepared for the communicative demands of their future academic and professional careers. However, in order to evaluate whether the Writing Project had achieved its aim of giving students opportunities to consolidate, develop and communicate their scientific understanding, we felt that it was necessary to systematically assess students' performance across the three reports.

The concern of this evaluation of the Writing Project was not whether students' writing had improved. We had no previous writing performance against which to compare their writing. Nor did we feel that improvement in writing could be measured across a period of six months. We were rather interested in facilitating a process whereby assessing students' writing enabled staff to critically reflect on the extent to which the curriculum (including the writing assignments) was enabling students to gain conceptual, methodological, and epistemological access to the discipline.

Of particular concern was the performance of students who come from educationally disadvantaged backgrounds, who make up approximately one half of the second year Chemistry class. These students, by their admission to UCT, have successfully hurdled the tremendous obstacles of apartheid education yet begin their university careers seriously disadvantaged compared to their white peers. Although most of these students benefit from intensive support in their junior courses, there is no illusion that by the second academic year of study the "gap" had been fully bridged.

My agenda-the Writing Centre's agenda-was a "transformative" one whereby the collaborative process "enables people to change by encouraging self-reflection and a deeper understanding of their particular situations" (Lather, 1986). Through this evaluation I was seeking to make visible, challenge and maybe shift deeply-rooted notions of writing and its role in the curriculum. Gee (1990) claims, "all practice (human action) is inherently caught up with usually tacit theories that empower or disempower people or groups of people." The teaching and assessing of writing is no exception.

NOTIONS OF WRITING: INSTRUMENTAL OR CONSTRUCTIVE?

The curriculum developments described above are a noteworthy achievement given that prior to the initiation of the Writing Project, this Chemistry department was typical of highly numerate scientific disciplines where writing skills are expected and assumed at senior levels but given very little, if any, curricular space for development in the undergraduate curriculum. There are many reasons for academics' resistance to the *explicit* development of writing within an undergraduate curriculum. Among them is a "survival of the fittest" mentality about the higher education enterprise where assessment in junior courses aims primarily to "weed out" weaker students. Attempts to initiate students into the mysterious demands of academic writing are perceived as "remedial activities" which compete with disciplinary content for precious curricular space.

Another reason for this resistance to the explicit teaching of writing in the discipline are deeply-rooted assumptions about what writing is and its role in the educational process. Moore (1996) discusses two contrasting views often held simultaneously by academics on the role of writing in student learning. The first view, what Moore calls the instrumental view, understands writing to be "merely a technical process of transmitting finished thought from mind to paper." Writing is reduced to a "generalizable skill, learned once and for all at an early age" (Russel, 1991, p. 295). In this view students write in order to display their mastery of content. Moore discusses three consequences of such an instrumental view of writing. One, students' writing problems are diagnosed as errors of syntax and grammar, which

can be “fixed” by closer attention to the appropriate rules. Two, students’ inability to write at the university is a failure of the schooling system and not a legitimate concern of academe. Three, students’ inability to write is a passing crisis of illiteracy to be responded to (if at all) with various forms of temporary remediation-what Rose calls the “myth of transience” (1985).

The second view of writing, the constructivist view, sees writing as a “much more complex business which consolidates-and even advances-thought and learning in unique ways” (Moore, 1996, p. 25). In this view, writing is writing about something. Understanding of the “something” and the business of writing are inseparable. In this view, students write, not merely to showcase their understanding but as central part of the process of *constructing* understanding. Crucial to this construction is a conceptual grasp of the disciplinary content.

Although the Writing Project was evidence that academics in the Chemistry Department valued and were willing to take responsibility for the development of writing in their discipline, there were signs that instrumental understandings of writing were fundamentally informing the way in which writing was being approached. Unlike their predecessors, Chemistry students were now given opportunities to write. But (as its name might suggest) the Writing Project was still seen as adjunct rather than intrinsic to the curriculum. Despite the careful design and mediation of the tasks, the writing assignments were delivered into a pedagogical and curricular vacuum.

Two examples will illustrate. In the Chlorine report students were required to discuss the environmental impact of a certain chemical industrial process. When I questioned whether the students had been exposed to the environmental debate in lectures, the lecturer responded that he did not have time to cover these debates in his lectures. Asking the students to discuss environmental issues in the assignment was a means of “covering more content.” Besides, he assumed most students already had an environmental awareness as part of their “cultural capital.” If this assumption was correct, then it might be fair enough to make such a discussion part of the assignment. If his assumption, was incorrect (I personally doubted his assumptions), then some students would be clearly disadvantaged in their report-writing before they ever began to write.

At the conclusion of one of the marking workshops, one staff member complained about the insipid quality of students’ conclusions, as though somewhere along the way the education system had failed to teach students to write conclusions. A more enlightened member of staff replied, “That’s because we don’t teach them to *make* conclusions!” I took her comment to mean that the undergraduate Chemistry curriculum did not initiate students into the processes of weighing evidence, in order to make even a tentative commitment to a particular position-surely a necessary component for drawing conclusions.

Because of deeply-rooted assumptions about writing and its role in the curriculum, some of the staff were unwilling (or unable) to concede that students’ ability to effectively communicate is inextricably bound to their conceptual and epistemological understanding of the discipline. Russell critiques American higher education writing-across-the curriculum initiatives in the 1990’s by saying, “until individual disciplines accept the responsibility of studying and teaching the writing of their community to students, writing-across-the curriculum will continue to be marginalized...writing will continue to be transparent and the myth of transience powerful among those who do not understand or acknowledge the relationship between writing and the creation and acquisition of knowledge” (1991, p. 298).

PORTFOLIO ASSESSMENT

In the U.S., U.K. and Australia there are a number of alternative methods of assessment being creatively applied to help teachers reflect on their students’ achievement in order to improve their own teaching practice. Portfolio assessment is one such method. One of the great strengths of portfolio assessment is the role it has played in strengthening what has historically been a tenuous link between the teaching of writing and its assessment. Judging a student’s writing ability on the basis of a portfolio of writing requires a look beyond simply the writing itself to the learning and teaching context for the writing. Thus, given the “textured and multi-layered” nature of portfolio assessment (Belanoff & Dickson, 1991), teachers recognize in portfolio assessment not only a valid measure of their students’ ability but also a valid measure of their own practice: Students’ portfolios can provide teachers with a unique opportunity to critically reflect on the effectiveness of their curriculum and teaching methodologies.

A survey of the most recent publications on portfolio assessment at college level (Belanoff & Dickson, 1991; Black et al., 1994) suggests that their most common uses are the assessment of individual student

performance for placement, into or exit from freshmen composition courses or other, senior writing courses. The use of portfolio assessment in the academic disciplines seems far less common (Lovitt & Young, 1994). And although portfolio assessment is increasingly being recognized as a powerful evaluative tool for staff and curriculum reform, in surveying the literature very few studies set out with this explicit evaluative agenda. Most users of portfolio assessment are mainly interested in assessing student performance. The implications of the assessment, for example for curriculum reform, are not the central aim of the assessment but an interesting-by-product.

Two reports in recent publications on portfolio assessment are exceptions. Larson (1991) describes a study at Lehman College of the City University of New York where portfolios were collected from a sample of students at various levels of the undergraduate programme in order to evaluate the effectiveness of a new curriculum on students' performance. In this instance, Larson states that they were less interested in the performance of individual students than they were in the performance of groups of students—students who had been taught with the new curriculum compared to those who had not. Although Larson's study highlights the potential of portfolio assessment for curriculum evaluation, like so much of the portfolio assessment literature, it is rich in descriptive detail of the portfolio process but leaves the reader no wiser about the effectiveness of the curriculum in question or the value of portfolio assessment as a means to **evaluate the curriculum**. Levitt and Young (1994) describe attempts at Clemson University to implement a portfolio assessment as part of a writing-intensive curriculum within the Department of Finance. The portfolios were meant not only for assessment but as "a basis for guiding the entire process of restructuring the curriculum." The strength of this contribution is the role that portfolio assessment can play in surfacing conflicting interests which exist between disciplinary academics and writing specialists. Both are interested in improving students' writing, but the latter intend to do so by "transforming the host culture." The host culture, of course, is not necessarily interested in being transformed! Since Levitt and Young's paper is work-in-progress, how the portfolio assessment actually served its intended purpose—that is, to guide the curriculum restructuring—is not dealt with in the paper. Thus despite Black's claim (1993) that, "portfolios, perhaps more than any other type of assessment tool, provide very rich information that can lead to the betterment of academic programs..." (p. 139), portfolio assessment for programme evaluation is relatively new and unresearched.

PORTFOLIO ASSESSMENT: THE PROCESS

Our particular assessment purpose (to gain an overall picture of student performance in order to evaluate the broader pedagogical and curricular context) as well as our particular context (an academic discipline) informed a number of choices that were made in designing our portfolio assessment project.

The Sample Size

Because we were interested not in students' individual performance but in their overall performance as an evaluation of the learning and teaching environment, only a representative sample of student writing needed to be assessed (White, 1994; CCCC Committee on Assessment, 1995). We decided that a random sample of approximately one third of the class would give us a representative performance picture. Fifty-four students (out of a total of 173) were randomly selected. In the end, five students' portfolios were withdrawn, leaving 49 students—a 28% sample size.

The Content of the Portfolio

Each student's portfolio contained the three final draft reports. To prevent a bias in the portfolio scoring, students' original reports were photocopied before they were submitted to their lecturers for the normal course assessment. All student identity marks were removed from the reports, and the reports were numbered. Name and number were recorded separately so that we could look at the performance of particular groups of students within the sample at a later stage.

The manner in which portfolios are assessed seems to vary almost as widely as their purposes. As portfolio assessment has gained in popularity a number of portfolio advocates are raising critical questions of holistic scoring—the most typical form of scoring in portfolio assessment. Some question the usefulness of a single score that only has meaning in a specified context. Even strong advocates of holistic scoring, like White, admit that one of the limitations of holistic scoring is that it has very little diagnostic power. Teachers, researchers and students are often interested in more than simple rankings—they want information they can use. Hamp-Lyons (1995) criticizes holistic scoring for “being a closed system, offering no windows through which teachers can look in and no access points through which researchers can enter” (p. 760).

More critically, some are questioning the validity of holistic scoring. Can such a complex phenomenon as writing be reduced to a single score? One of the strongest critics, Elbow (in Hamp-Lyons, 1995) argues that the “simple-minded, single-number holistic score” reduces writing to “some-thing which can be scored on a scale that is not only quantitative but one-dimensional” (p. 452). Advocates of holistic scoring argue that, rather than “reducing” the complex phenomenon of writing, holistic scoring recognizes that writing ability is more than the sum of its parts. Attempts to break writing into its “parts” (conventional subskills include, e.g., organization, content, spelling, mechanics) can cause frustration for markers and result in unreliable outcomes since there is very little professional consensus on the definition and importance of these subskills (White, 1994).

At the time we designed our scoring guide, I was concerned about what I perceived to be the limitations of holistic scoring for our purposes. My concerns were: First, our assessment framework needed to serve our “transformative agenda”—that is, enable the staff to critically reflect on their pedagogical and curricular practice. Second, it needed to be true to our views about writing—that is, it needed to reflect our theoretical understanding of writing as a constructive act. Third it needed to reflect the reality of our context—that is, the teaching of writing in the context of a discipline. These purposes are interrelated strands but will be addressed separately.

First, our assessment framework needed to serve our “transformative agenda.” If the portfolio results were to enable staff to reflect critically on the pedagogical and curricular environment, then we needed diagnostic information. Accepting the problematic nature of these ill-defined writing “subskills” (which holistic scoring advocates critique), our assessment purpose required some indication of where students’ problems lie and therefore where curricular adjustments might be needed.

Second, our assessment framework needed to reflect our theory of writing. Constructivist theories of writing suggest, as stated earlier in the paper, that writing is a part of the process of constructing understanding, and crucial to that construction in an academic context is a conceptual grasp of the disciplinary content. We therefore needed criteria that reflected the interdependence between students’ conceptual understanding of the scientific concepts and their ability to communicate that understanding effectively. Our theory could not support a scoring guide where “language skills” were assessed separate from “content understanding,” as is often common practice in assessing scientific writing. These dichotomies seem to reflect instrumentalist views of writing.

Third, our assessment framework needed to reflect the reality of our context—that is, a department where academics would readily admit their inexperience in teaching, assessing, even talking about student writing. This meant that both the process of designing the scoring guide, as well as the final product, needed to be “owned” by the Chemistry staff. Whatever criteria or guidelines were used for judging good versus poor writing, they needed to be authentic to the context of their discipline, “as close to the heart of a genre or problem as possible” (Wiggins, 1994, p. 139). Hamp-Lyons argues that methods of scoring need to be “constructions of *local theory* of what good writing is” (1995, p. 454). This seems particularly important for the teaching and assessing of writing in disciplines that may hold very different values than those of traditional Arts or Social Science disciplines.

We decided therefore that for our purposes it was crucial that the portfolios be assessed according to an explicitly-stated set of criteria. Although undoubtedly Chemistry lecturers have an intuitive understanding (rather than a coherent theory) of what good writing is, the explicit criteria for good writing are seldom articulated amongst colleagues (let alone to students!). A workshop was therefore set up with the goal of identifying a set of criteria for scientific writing ability at this level. In the workshop the Chemistry staff members were each given three anonymous student scripts selected from the first writing assignment. Their task was to score the scripts on a 1-3 scale and to be prepared to justify their scores.

The workshop proved to be an important forum for a “curricular and pedagogical dialogue” between writing and discipline experts which Jones and Comprone (1993, p. 61) argue is essential for effective writing-across-the-curriculum initiatives. Senior academics with years of lecturing and research experience in the discipline wrestled with the task of arguing why paper A was better (or worse) than paper B. The process of justifying their score was one which opened up for scrutiny the tacit set of “choices, guesses and assumptions” (Gee, 1994, p. 84) which lead to a marker’s particular interpretation and subsequent judgement of a student’s writing. For example, these may be assumptions about what prior knowledge students bring with them; guesses at students’ meaning-making attempts when their choice of linguistic tools is limited; *choices* about whether to value grammatical accuracy over conceptual risk-taking which can often lead to syntactic breakdown. The workshop forced these “choices, guesses and assumptions” into the open where they could be examined and challenged. Lecturers’ expectations or “standards” emerged from the workshop with new-found clarity—a process which empowered not only the lecturers but students as well.

The result of this lively debate was a set of five criteria for assessing the three scientific reports (see Appendix 1 for full details):

- has the student clearly understood the task/problem?
- does the student understand the science involved?
- has the student constructed a logical argument?
- has the student reached a conclusion?
- have the scientific conventions of the task been followed?

According to this set of criteria, a skilled writer is not simply one who has accurately followed the rules that govern the English language and adhered to the specific conventions of scientific writing, as important as these are. In the Chemistry staff’s own words, a skilled writer is one who “reveals an understanding of the complexities of the problem,” “shows evidence of multiple perspectives,” “understands how the science is relevant to the problem,” “substantiates and exemplifies, not just states,” “constructs, not just cuts and pastes” (direct quotes from the workshop). This group of Chemistry staff’s local theory of writing was certainly showing strong hints of constructivist thinking. Perhaps in some small way notions of writing were shifting.

I was reluctant to let go of what I perceived to be the value of a holistic score; therefore, in the end our assessment framework gave markers opportunity to assess students’ reports according to the five criteria as well as holistically on a 1-4 scale (1 = weak, 2 = fair, 3 = competent, 4 = strong). Staff were given the choice of scoring the reports holistically first, and then scoring according to the criterion, or vice versa. Unfortunately markers’ assessment processes were not recorded; therefore, there was no means of tapping the relationship between the scores. For some, the holistic score may have been a “sum of the parts,” for others a “weighting of the parts” (Condon & Hamp-Lyons, 1994).³

The Assessment Procedure

The team of markers consisted of five members of the Chemistry Department and one member of the Writing Centre staff with a Chemistry background. I played the role of administrator and moderator, since as a non-scientist my scoring would not be reliable.

Our assessment procedure deviated from standard portfolio assessment practice in that we did not attempt to assign a single score reflecting the overall quality of the portfolio. Since we wanted diagnostic information, each report in a student’s portfolio was scored separately. Some might argue justifiably that we lost the “motion picture” nature of portfolio assessment. At best we had a panoramic view—a view of three “snap- shots”. However, what we lost in “over-all-view” we gained in diagnostic detail. Thus each report had an overall (holistic) score and a score for each criterion.

To ensure some measure of interrater reliability, a marker coordination workshop was held the day before the actual assessment. For the actual portfolio assessment each report was scored twice. Where there was a discrepancy of more than one point on the overall score, a third score was given for resolution. The logistics of resolving each criterion score seemed too great; it was decided to simply record the scores and determine interrater reliability when the data was analysed.

Interrater reliability was calculated in terms of the consistency between markers within one point. All scores had high levels of reliability: overall score (OS): .93, task (TA): .92, Science (SC): X3, Argument (AR): .87, Conclusion (CON): .82, Conventions (COV): .91. It is interesting to note the criteria on which markers attained the highest and lowest levels of agreement. It is rewarding to note the very high levels of agreements on the task criterion. The markers had been carefully taken through the task requirements and

therefore easily recognized “task understanding” (or lack of). On the other hand, there was less clarity when it came to assessing whether students had constructed a logical argument or reached a conclusion. What an argument with an appropriate conclusion consisted of in this context had never been explicitly discussed with the students or even among the lecturers themselves. Hence, it is not surprising that there was less agreement between markers on the argument and conclusion criteria.

Table 1. A Comparison of the Mean Overall Scores (OS) per Report (CA, CH, OC) per Group (EA, ED)

	CA-OS	CH-OS	OC-OS
EA	2.82	2.78	2.82
ED	2.07	2.03	2.26
	(<i>p</i> = .0007)	(<i>p</i> = .0003)	(<i>p</i> = .01)

The Assessment Results and Analysis

Once all the scores for the reports were collected, the data was analysed in four ways. First, the average overall score per report was 2.4. Thus, the overall performance of students’ writing was assessed between “fair” and “competent” – not a surprising or particularly telling result.

The second analysis was a comparison of the overall scores of various groups of students within the total sample, that is, male/female, Science/ Engineering students, and educationally advantaged (EA) educationally disadvantaged (ED) students. No significant difference was found between the scores of male and female students or the Science and Engineering students. But the comparison of the EA group (n = 23) with the ED group (n = 26) revealed a significant difference between the performance of the two groups on each of the three reports (*p* < .05). For each report the EA’s scores averaged near “competent”, and the ED’s scores averaged “fair”.

The third analysis was a comparison of the overall scores (t-test: two sample assuming unequal variances) for each of the report types (i.e. Caffeine, Chlorine and Oxalyl Chlorine). As indicated in Table 1, the EA group’s average overall scores were 2.8, 2.7 and 2.8 and the ED group’s average overall scores were 2.0, 2.0 and 2.2, respectively. There was no significant difference between the average overall score of the reports for either the EA or ED group, that is, both groups’ performance across the three reports was consistent.

This is a surprising result since research into the role of task suggests that due to a range of variables (e.g., students’ familiarity with the genre, the conceptual complexity of the task, the clarity of the task design, the forms of task mediation, students’ motivation for the task) students may perform unevenly on different tasks (Inglis & Yeld, 1994). Given the conceptual complexity of the Oxalyl Chloride report, in addition to the unfamiliar genre, I had anticipated the possibility of declining scores relative to the less demanding requirements of Caffeine report. One possible explanation for this surprising result is that although we would expect students’ performance to fluctuate depending on task complexity, familiarity with gem-e, etc., the more complex tasks were appropriately sequenced and mediated so that students were gradually acquiring the subject matter expertise needed to handle more conceptually challenging tasks. This of course would be the most desirable explanation.

There are other possible explanations, however. Even though the scoring was meant to be criterion-referenced, in assigning a holistic score to the report the markers were unconsciously norm-referencing on each task that is, assessing the report relative to the other reports rather than based on actual achievement on that particular assignment. The result is a normal distribution of overall scores for each report with a very similar mean score.

Table 2. A Comparison of Criterion Scores (TA, SC, ARG, CON, COV) per Report (CA, CH, OC) for Educationally Advantaged (EA) Group

	CA	CH	OC
TA	3	2.8	3.1
SC	2.6	2.6	2.7
ARG	2.6	2.6	2.7
CON	2.6	2.6	2.8
COV	2.7	2.8	2.9

Upon closer examination of individual students' overall scores (as opposed to merely an analysis of the mean overall score for the group), there is fluctuation of performance, although not in any consistent pattern. Ten out of the 23 EA students' scores varied by more than one across the three reports (e.g., for the Caffeine, Chlorine and Oxalyl Chloride reports, scores of 1,3,2 respectively; 4,2,4 respectively; 1,2,3, respectively). However, only 3 out of 26 ED students' scores varied by more than one point across the three reports. The lack of fluctuation in individual ED overall scores across the three reports may be the result of poor language and conceptual skills, placing a "ceiling" on their performance.

The fourth analysis was a comparison of the criterion scores for each report type. It is interesting to note that the average task (TA) criterion scores are higher than the other criterion scores for both groups (see Tables 2 and 3), and for the ED group the science criterion scores tend to be lower than the other criterion (see Table 3). However a comparison of mean (ANOVA) revealed no significant difference between the criterion scores for each of the reports for either the EA or ED group; that is, the scores on the individual criteria were consistent across reports for both groups.

These results are also somewhat surprising and disappointing given the diagnostic purposes for the exercise. Again this could be because of the unconscious norm-referencing of the markers. Another speculation for this result is that the criteria themselves do not in fact reflect independent skills or abilities. Perhaps White (1994) is right when he argues as noted earlier that attempts to break writing into "parts" can result in unreliable outcomes. Another speculation is that a four point scale is too crude for our purposes. Had the markers been assessing on an 8 or 10 point scale perhaps there would have been greater variance in the scoring (interrater reliability would have been much lower, however).

Table 3. A Comparison of Criterion Scores (TA, SC, ARG, CON, COV) per Report (CA, CH, OC) for Educationally Disadvantaged (ED) Group

	CA	CH	OC
TA	2.3	2.2	2.5
SC	1.9	1.9	1.8
ARG	2.0	1.8	2.0
CON	1.9	2.0	2.3
COV	2.0	2.0	2.3

DISCUSSION

The portfolio assessment results did not prove anything, nor were they meant to. The desired response from the exercise was that staffs understandings of writing and its role in the curriculum would be challenged, maybe even shifted as a result of the whole portfolio exercise, in particular the design of the criteria and their systematic application to student writing. Nonetheless the quantitative results which emerged proved to be powerful, albeit tentative indicators of underlying curricular problems which prior to the portfolio assessment had been unacknowledged.

The lower overall scores of the ED group suggest they are struggling to effectively communicate in report-writing. This finding in itself did not surprise staff since it is accepted that these students have come

from an educational system which was “systematically structured to underequip black students for higher education” (Startield, 1994, p. 176). The data do not shed any light on why the ED students are struggling. Nonetheless it is crucial that we explore possible interpretations for the data in order to gain more insights for restructuring the curricula to meet the learning needs of the soon-to-be majority at UCT.

One speculation for these results is that the ED’s relatively weak reports are a result of poor language skills. Underlying this analysis is the assumption that “at university level, students will not be able to make meaning or construct the relevant knowledge if their language ability in the medium of instruction is weak” (Inglis & Yeld, 1994). That is, in the case of the ED group, they are struggling to produce “competent” reports because they have to understand, read and write in English. As noted earlier, this is the most common interpretation advanced by lecturers in their assessing of the writing of second language or nonstandard dialect speakers of English.

Another speculation, however, is that the weak reports are a result of the students’ inability to make sense of what they are writing about. Inglis and Yeld (1994) found in their research on assessing English second language students’ scientific writing that one student’s English on a particular task may differ markedly from that same student’s written English on another task, *if the conceptual demands of the tasks differ markedly*. They conclude that poor control of written English may indicate poor grasp of the material required for that task, and *not* what is often inferred: namely, that the reason why the student has not understood the material is because of poor control of written English.

Whether limited language skills is blocking a student’s access to the conceptual understanding or whether a poor conceptual understanding is blocking a student’s full use of their language skills seems to be one of those proverbial chicken-and-eggs debate. The portfolio results give no clear evidence in support of one speculation over the other. In fact, I would argue that as long as one views writing as inseparable from what one is writing about, then either interpretation is plausible. In fact, it is most likely that both are happening at the same time. They are two sides of the same coin since “learning is mediated through language” (CCCC Committee on Assessment, 1995, p. 434).

This is indeed the challenge for students acquiring academic literacy, in this case chemistry literacy, a form of literacy which “is integrally bound up with the systematic (but abstract) ways of thinking of (the) discipline” (Moore, 1996, p. 26). “A student is only able to articulate the discourse of a discipline once he/she has a conceptual grip of the material under discussion, and a sense of how language is used to express ideas in that discipline (Taylor in Moore, 1996). Moore discusses the curricular consequences of an ‘academic literacy’ view of writing—where writing cannot be understood separately from what is being written about. The first consequence is that greater prominence be given to the role of writing in reflecting and advancing the conceptual learning goals of a discipline. The second consequence is a realization that the curriculum needs to pay greater attention to the opportunities it provides for students to explicitly grapple with the language of the discipline. Moore (1996) writes, “this includes active approaches to reading, listening and discussion of concepts, an apprenticeship to both the concepts *and the language by which they are constituted*” (p. 27).

OUTCOMES OF THE PORTFOLIO ASSESSMENT PROCESS

The aim of the portfolio assessment was, through an assessment of students’ writing, to encourage staff to critically reflect on the extent to which the curriculum is enabling students to gain conceptual, methodological, and epistemological access to the discipline. Ultimately this critical reflection should result in the restructuring of curricula which are more sensitive to the needs of a diverse student population.

The results of the portfolio assessment, particularly the performance of the ED students, suggest to the Chemistry staff that the learning needs of students represented by the ED group are not being catered for in the second year Chemistry curriculum. Students producing on average “fair” reports are unlikely to be adequately prepared for the demands of their future academic and professional choices. The issues of dealing with tremendous diversity of preparedness in the same classroom are complex, and the Academic Development has only begun to explore the range of complex problems and their curricular implications. The immediate question that needed to be addressed was: what were the implications of the portfolio assessment project for the planning of next year’s course?

At the end of the year, the evaluation report (written by me and endorsed by the team) made the

- i. Given the ED students' performance on the reports, the second year teaching staff should review the relationship between the lecture content, the practicals and the writing assignments. The aim of the writing assignments should be for students to consolidate and develop their understanding of material presented in lectures, practicals and textbooks.
- ii. Given the considerable effort required to sustain developmental initiatives of this kind, the Chemistry Department and the Science Faculty as a whole should consider ways to value and reward staff engaged in these efforts.

Underlying these recommendations lies the assumption that improving students' writing is fundamentally a curricular and pedagogical issue. As Moore (1996) concludes, "Successful writing...seems to be one of the most important outcomes of the entire curricular process" (p. 82). The Writing Project needed to aim beyond simply getting students to write, even getting students to write in genres appropriate to the discipline. The aim should be that, through writing, students are taken into the conceptual, analytical and epistemological heart of the discipline. The result-students will be not only better critical thinkers but better communicators of that critical thinking.

It is heartening that the department approved the recommendations unanimously, opening an opportunity for potentially radical curriculum review. At the time of the writing of this article there have been significant developments.

1. In 1996 a member of the Chemistry staff was appointed to oversee the Writing Project, changing the Writing Centre's role from "driver" of the Writing Project to "consultant."
2. This year's writing assignments were very successfully developed with a deliberate attempt to integrate the writing assignments with the lecture content.
3. A number of measures were put in place to ensure that students' diverse levels of needs were catered for:
 - writing workshops prepared students for the written assignments by providing them with opportunities to explicitly develop chemistry literacy through exposure to, and discussion and analysis of a range of scientific texts.
 - the process approach to writing was strongly encouraged, and students were provided with intensive feedback while drafting.
4. The portfolio assessment project has had a significant impact on the day-to-day assessment practices within the course. For example, assessment frameworks are now designed for each assignment and given to students as well as markers; marking workshops for lecturers are standard procedure now. All of these are very encouraging steps.

The extent to which these steps will lead to long-term transformation remains to be seen. The unanimous approval of the second recommendation that the department consider ways to value and reward staff engaged in developmental initiatives-remains an approval of ink-on- paper only. Even if some staff are willing to invest in such initiatives, will curricular innovation survive in a culture that "devalues undergraduate education in favor of professional research"? (Anson & Brown, 1991, p. 257). In the context of an institution like UCT, despite current pressures for political and social transformation, the curricular and pedagogical implications of portfolio assessment are in conflict with personal, departmental and even institutional interests. Anson and Brown argue that "contemporary writing theory is radically disruptive to some of the key ideological premises of research universities" (p. 263), and sensible academics are understandably reluctant to work against their own interests.

CONCLUSIONS: PORTFOLIO ASSESSMENT AS A POTENTIAL CATALYST FOR STAFF AND CURRICULAR REFORM

The potential of the portfolio assessment process as a catalyst for critical reflection about writing is indisputable. The literature on portfolio assessment is rich in testimonies of how the portfolio experience can change teachers' understandings of writing and its role in the curriculum. Our project was no exception. One member of the Chemistry Department who participated in the portfolio assessment reflected,

"I'd not thought much about writing skills before last year (the initiation of the Writing Project). I assumed that students learnt these "along the way." But reading students' answers to test questions (as well as their writing assignments) has convinced me that a student's ability to explain a scientific concept is a very good measure of that student's understanding...but it's also clear that we can't expect students to learn effective communication skills without training."

Although a view of writing as the act of constructing understanding (as opposed to merely "a good measure" of that understanding) is not fully articulated here, her comment reflects an evolving understanding of writing-that writing skills are not just learned "along the way" but need to be imparted to students in explicit ways. These evolving understandings are crucial, yet ultimately they must result in the restructuring of curricula more sensitive to the needs of a diverse student population.

Our experience with portfolio assessment suggests that its potential as catalyst for curriculum reform rests on a number of issues.

As was the case in the Chemistry portfolio project, it is essential to look at the performance of groups of students within the larger sample with a particular focus on students from historically marginalized sectors. To simply report the overall average scores as an indication of the success of any programme is highly deceptive.

Our assessment framework was of limited diagnostic value, and only further research will shed light on whether the problem lies in the definition of the criteria and/or in the range of the scale. Nonetheless, I would maintain that the design and application of an agreed-upon set of criteria to student writing played an essential staff development role and ultimately validated the whole project. I would argue with Gipps (1994) that "group moderations...is the most thorough of the quality assurance approaches" (p. 142).

Although our assessment framework encouraged markers to record any qualitative assessments of students' writing (in addition to their quantitative assessments), in the end very little attention was devoted to these. Future assessment processes need to devote much more time to discussing lecturers' interpretations of the scores and the curriculum implications of these interpretations. Black (1993) argues that the unique potential of portfolio assessment for program evaluation will lie in its ability to merge both quantitative and qualitative data. We clearly have only scratched the surface.

Amiran and Courts (1994) refer to the development of a portfolio scoring system for program assessment as a "spiral process" where one moves from curricular or program aims to an analysis of the portfolios and back again, from discussion of particular portfolios' characteristics to criteria and back again, over a period of years. Clearly, our spiral has just begun. Nonetheless, at this point in the spiral, portfolio assessment opened up a unique space in this department for dialogue, dialogue which reflects, if only dimly, new understandings about writing and its role in the curriculum. But the ultimate aim is that students-particularly those who for too long have been excluded from the dialogue-will graduate as confident participants in their own dialogues, dialogues that matter "out there."

NOTES

1. The term 'black' has been used traditionally to refer to all who were to a greater or lesser extent discriminated against by apartheid legislation, that is, in this case, African, Coloured (people of mixed race) and Indian students.
2. I acknowledge that all forms of labelling are contentious and this is nowhere more true than in South Africa. The term "educationally disadvantaged" usually refers to African and Coloured students who completed their schooling in what was until 1996 a racially divided, highly under-resourced educational system. Their disadvantage is compounded by the fact that the majority of these students are speakers of English as an additional language. Conversely, "educationally advantaged" students refer to white students with either English or Afrikaans as a home language from what were relatively very well-resourced and privileged schools.
3. See Condon and Hamp-Lyons' research (1994). In tracing readers' assessment processes they contrast "impression of the whole" holistic judgement with "weighing of the parts" judgements.
4. Students from both of these faculties are required to do the Chemistry course. This analysis was done as there was some speculation that Engineering students are better equipped to write than Science students.
5. Although our two groups match the profiles described in note 2, we must be cautious of over-simplifying diversity. Even amongst our ED group there was tremendous diversity of "preparedness" both in terms of the quality of schooling in the ex-African and Coloured Education Departments as well as their exposure to English; and from students who are virtually bilingual to students whose only exposure to English was with teachers who may or may not have been competent English speakers themselves. Furthermore, major transformation in the education system will, hopefully radically change these historical profiles so that language background (and hence racial classifications) will no longer be synonymous with levels or preparedness.

REFERENCES

- Allie, S., Kaunda, L., & Holmes, P. (1995). *Communication skills in the undergraduate curriculum in Science*. Unpublished report to the Committee on Undergraduate Education in Science, University of Cape Town, South Africa.
- Amiran, M., & Courts, P. (1994). *A stab at scoring portfolios*. *Assessment Update: Progress, Trends, and Practices in Higher Education*, 6(2), 6-7.
- Anson, C., & Brown, R. (1991). Large-scale portfolio assessment: Ideological sensitivity and institutional change. In P. Belanoff & M. Dickson (Eds.), *Portfolios: Process and Product*. Portsmouth, N.H.: Boynton/Cook Publishers.
- Belanoff, P., & Dickson, M. (1991). Introduction. In P. Belanoff & M. Dickson (Eds.) *Portfolios: Process and Product*. (pp. xix-xxiv). Portsmouth, N.H.: Boynton/Cook.
- Black, L. (1993) Portfolio assessment. In T. Banta & Associates (Eds.), *Making a Difference: Outcomes of a decade of assessment in higher education*. San Francisco: Jossey-Bass.
- Black, L., Daiker, D., Sommers, J., & Stygall, G. (1994). New directions in portfolio assessment: Reflective practice, critical theory, and large scale scoring. Portsmouth, NH: Heinemann-Boynton/Cook.
- College Composition and Communication Conference Committee on Assessment (1995). Writing assessment: A position statement. *College Composition and Communication*, 46(3), 43W37.
- Condon W., & Hamp-Lyons, L. (1994). Maintaining a portfolio-based writing assessment: Research that informs program development. In L. Black, D. Daiker, J. Sommers, G. Stygall (Eds.), *New directions in portfolio assessment: Reflective practice, critical theory, and large scale scoring*. (pp. 277-285). Gee, J. (1990). *Social linguistics and literacies: Ideology in discourse*. London: The Falmer Press. Gipps, C. (1994). *Beyond testing: Towards a theory of educational testing*. Washington, DC: The Falmer Press.
- Hamp-Lyons, L., & Condon, W. (1993). Questioning assumptions about portfolio-based assessment. *College Composition and Communication*, 44(2), 176-190. Hamp-Lyons, L., (1995). Rating nonnative writing: The trouble with holistic scoring. *Teaching English to Speakers of Other Languages Quarterly*, 29(4), 759-762.
- Hamp-Lyons, L., (1995). Uncovering possibilities for a constructivist paradigm for writing assessment. *College Composition and Communication*, 46(3), 446-455.
- Jacobs, D., Moore, R. (1996). *Concept-driven teaching and assessment in the Zoology invertebrate course at UCT*. Paper presented at 4th annual meeting of the Southern African Association for Research in Mathematics and Science Education, Peterborough, South Africa.
- Inglis, M., & Yeld, N. (1994) *Input vs. attribute-based language proficiency*. Paper presented at 2nd annual meeting of the Southern African Association for Research in Mathematics and Science Education, Durban, South Africa.
- Jones, R., & Comprone, J. (1993). *Where do we go next in writing across the curriculum?* *College Composition and Communication*, 44(1), 59-68.
- Larson, R. (1991). Using portfolios in the assessment of writing in the academic disciplines. In P. Belanoff & M. Dickson (Eds.), *Portfolios: Process and Product*. Portsmouth, N.H.: Boynton/Cook Publishers.
- Lather, P. (1986). Research as praxis. *Harvard Educational Review*, 56(257-277).
- Lovitt, C., & Young, A. (1994). Portfolios in the disciplines: Sharing knowledge in the contact zone. In L. Black, D. Daiker J. Sommers, G. Stygall (Eds.), *New directions in portfolio assessment: Reflective practice, critical theory, and large scale scoring*. (pp. 334-346). Portsmouth, NH: Heinemann-Boynton/Cook.
- Moore, R. (1996). *The role of student writing in learning Zoology*. Unpublished M.Ed. thesis, University of Rhodes, Grahamstown, S. Africa.
- Moore, R., Paxton, M., & Shay, S. (1996). *Review and evaluation of writing across the curriculum initiatives at a South African university*. Unpublished paper, University of Cape Town, S. Africa.
- Rose, M. (1985). The language of exclusion: Writing instruction at the university. *College English*, 47(4), 341-359.
- Russell, D. (1991). *Writing in the academic disciplines, 187&1990: A curricular history*. Carbondale, IL: Southern Illinois University Press.
- Starfield, S. (1994). Cummins, EAP, and academic literacy. *Teaching English to Speakers of Other Languages*, 28(1), 176-179.
- Stygall, G., Black L., Daiker D. & Sommers, J. Introduction (pp. 1-9). In L. Black, D. Daiker, J. Sommers, & G. Stygall (Eds.), *New directions in portfolio assessment: Reflective practice, critical theory, and large scale scoring*. Portsmouth, NH: Heinemann-Boynton/Cook.
- Taylor, G. (1998). *Literacy by degrees*. Milton Keynes: Open University Press.
- White, E. (1994). *Teaching and assessing writing*. San Francisco: Jossey-Bass Publishers.
- Wiggins, G. (1994). The constant danger of sacrificing validity to reliability: Making writing assessment serve writers. *Assessing Writing*, 1(1), 129-139.
- Yeld, N. (1992). *Discussion paper: Writing Programmes at University of Cape Town*. Unpublished paper: Academic Development Programme, University of Cape Town, South Africa