

NQF SUPPORT LINK
READER FOR MODULE 4

Acknowledgements

We would like to acknowledge the use of the following extracts and thank the publishers and authors for making them available:

Module 1: Implementing the NQF

Van Schaik Publishers, Pretoria: Coetzee, M (2002) Getting your Accreditation. The quality assurance and assessment guide for education, training and development providers.

Module 2: The NQF and Strategic Governance

Reading 3: De Clerq, F. (2002). Decentralisation of Authority to Districts: Search for District development and/or Control? Wits School of Education, Johannesburg (Presented at the International Conference on Education and Decentralisation, African Experiences and Comparative Analysis).

Reading 4: Maclennan, A. C., Education Governance and Management in South Africa. Unpublished Doctoral Thesis, University of the Witwatersrand.

Reading 5: Day, Professor C. (1999) Professional Development and Reflective Practice: purposes, processes and partnerships. In Pedagogy, *Culture & Society* 7 (2), pp. 221-233.

Toolkit 4 and Toolkit 6: Pearson Education, Edinburgh: Adapted from Koch, R. (2000) The Financial Guide to Strategy: How to create and deliver a useful strategy.

Module 3: The NQF and Learning Programmes

Reading 2: Juta Academic Publishing: "Definitions of Curriculum" from Angelis, D., & Marock, C (2001). Curriculum implication for FET. In Angelis, D., Lolwana, P., Marock, C., Matlhaela, P., Mercorio, G., Tsolo, S. & Xulu, S. The further education and training institutional readiness handbook: making learning work. Toolkit 8 – Listening Techniques Listening Exercise from Hope, A. & Timmel, S. (1988). Community Workers' Handbook 2. South Africa: The Grail, Lumko Missiological Institute and the Federation of Dominicans of Southern Africa (FEDOSA).

Module 4: The NQF and Assessment

Readings 1-4: LGWSETA Code of Conduct

Reading 5: Falmer Press: Gipps, C. (1994) Beyond Testing, pp3-4.

Reading 6: Slippery Rock University: Governing Principles for Assessment of Student Learning.

Reading 7: Vivlia Publishers: Kramer, D. (1999) OBE Teaching Toolbox, pp. 58-63,

Reading 8: Pahad, M. Approaches to Assessment and Evaluation. Wits/IEB Coursebook, pp. 34-37.

Reading 9: Bulman, F. Peer Assessment in Groups. Portland State University Center for Academic Excellence (web site)

Module 5: The NQF and Learnerships

Readings 27-30: Knowledge Resources: Hattingh, S. Extracts from the ROADMAP Series to Learning and Skills Development, pp. 12-13.

Reading 31: Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH: Extracts from Heitmann, W. "The action-oriented learning approach for promoting exemplary job performance and employability" (Unpublished article).

Readings CS 1-CS35: Charlene d'Hotman, Unpublished case study, Central Johannesburg College and Tshwane North College.

Reading 1



CODE OF CONDUCT TO BE SIGNED BY REGISTERED ASSESSORS

Registered Assessor Number : 1:2002

I, the undersigned, am applying to become a registered LGWSETA assessor. I agree that, if my application is successful, I hereby commit myself to abide by the LGWSETA Code of Conduct in relation to all my work conducted as an LGWSETA assessor. The Code of Conduct to which I agree is as follows:

- I shall conduct my work as an LGWSETA assessor with integrity, seeking at all times to create a positive environment for assessment and to take note of and respect the historical diversity of candidates' cultural, linguistic and educational backgrounds;
- any conflict of interest such as a financial or family relationship or close friendship existing between myself and any candidate shall be declared in advance, and, if requested, I shall recuse myself from the process in such instances;
- all information received during my work as an assessor about individuals or organisations will be treated with the strictest confidentiality unless it is relevant to the fairness, reliability and validity of the assessment process;
- all relevant information about any irregularities in the assessment process of which I become aware will be included in my reports to the moderator; these will include:
 - unplanned environmental, personal or other problems which may have interfered with the performance of the candidate
 - suspected or proven irregularities committed by the candidate
 - suspected or proven irregularities committed by any other parties to the assessment
 - any suspected or proven bribery, threats or sexual or other harassment of or by candidates
 - any grounds for doubting the authenticity of the evidence presented during the assessment process;
 - if I have reason to believe the moderator is not addressing irregularities brought by myself to his or her notice, I shall **draw these** irregularities to the attention of the LGWSETA ETQA Manager;
 - any constructive comments about the standards or qualifications which I am assessing will be included in my reports to the moderator, to be collated and forwarded to the LGWSETA; these will be forwarded to the relevant SGB for inclusion in the SAQA review process.

NQF SUPPORT LINK MODULE 4

I shall conduct my work in line with the vision and mission of the LGWSETA, particularly in relation to improving the quality of education, training and assessment for learners in the sector, giving guidance and support to all learners to achieve their full potential.

Signed: _____ Date: _____

Name in full (printed): _____

Reading 2

Code of Conduct

Registered Assessors

As a registered Assessor of the XXX SETA, I will carry out Assessments which are fair, valid and reliable.

I agree to abide by the principles of Transparency and Confidentiality

I will conduct assessments and perform the duties Of an assessor With the interests of the learner at heart and Without compromising the integrity of the XXX SETA and the Assessment system in any way

I will use the assessment criteria and the assessment guides For each unit standard as the ultimate measure of the Learner's Competence in that unit standard

I will adhere to all the requirements Of the XXXSETA in promoting The quality of education and training and assessment For all learners, giving guidance and support To all learners to achieve their full potential.

Signed

Witness

Date

Date

XXXXSETA

This registration as an assessor is valid for a period of three (3) years

Reading 3



CODE OF CONDUCT TO BE SIGNED BY REGISTERED MODERATORS

Registered Moderator Number : 1:2002

I, the undersigned, am applying to become a registered LGWSETA moderator. I agree that, if my application is successful, I hereby commit myself to abide by the LGWSETA Code of Conduct in relation to all my work conducted as an LGWSETA moderator. The Code of Conduct to which I agree is as follows:

- I shall conduct my work as an LGWSETA moderator with integrity, seeking at all times to create a positive environment for assessment and moderation and to take note of and respect the historical diversity of candidates' and assessors' cultural, linguistic and educational backgrounds;
- any conflict of interest such as a financial or family relationship between myself and any candidate or assessor shall be declared in advance, and, if requested, I shall recuse myself from the process in specific instances;
- all information received during my work as a moderator about individuals or organisations will be treated with the strictest confidentiality unless it is relevant to the fairness, reliability and validity of the assessment process;
- any unevenness in the standards of different assessors will be noted, and every effort will be made through feedback and constructive support to achieve a common standard amongst all assessors under my moderation; in this way I shall try to promote quality assessments and avoid standards drift;
- if I identify a particular assessor as unreliable and am unable to correct his or her weaknesses through a supportive and transparent process I shall inform the LGWSETA ETQA about this problem and ask them to take appropriate steps;
- all relevant information about any irregularities in the assessment process of which I become aware will be included in my reports to the LGWSETA ETQA; these will include:
 - unplanned environmental, personal or other problems which may have interfered with the performance of the candidate
 - suspected or proven irregularities committed by the candidate
 - suspected or proven irregularities committed by any other parties to the assessment
 - any suspected or proven bribery, threats or sexual or other harassment of or by candidates or assessors
 - any grounds for doubting the authenticity of the evidence presented during the assessment process;
 - any constructive comments about the standards or qualifications which I receive from assessors or note in my own work shall be collated and

included in my reports to the LGWSETA ETQA Manager; these will be forwarded to the relevant SGB for inclusion in the review process;

- any constructive comments about the assessment guides, instruments or procedures which I can make by reflecting on their use on the ground shall be collated and sent to the LGWSETA ETQA Manager; this feedback will be used to improve old guides and instruments and develop new ones in a continuing process of renewal and improvement of quality.

I shall conduct my work in line with the vision and mission of the LGWSETA, particularly in relation to improving the quality of education, training and assessment for all learners and assessors in the sector, giving guidance and support to all involved to achieve their full potential.

Signed: _____ Date: _____

Name in full (printed): _____

Reading 4



CODE OF CONDUCT TO BE SIGNED BY REGISTERED VERIFIERS

Registered Verifier Number : 1:2002

I, the undersigned, am applying to become a registered LGWSETA verifier. I agree that, if my application is successful, I hereby commit myself to abide by the LGWSETA Code of Conduct in relation to all my work conducted as an LGWSETA verifier. The Code of Conduct to which I agree is as follows:

- as an LGWSETA verifier, I shall seek at all times to ensure that assessments in the sector are fair, reliable and valid; that assessment processes are transparent, and, as far as possible without damaging their integrity, developmental and supportive to the learners.
- any conflict of interest such as a family relationship, close friendship or financial agreement between myself and any providers shall be declared in advance, and, if requested, I shall recuse myself from the verification process in such instances;
- all information received during my work as a verifier about individuals or organisations will be treated with the strictest confidentiality unless it is relevant to the fairness, reliability and validity of the assessment process or the quality of the provision of education and training;
- all relevant information about any irregularities in the assessment process of which I become aware will be included in my reports to the LGWSETA ETQA; these will include:
 - unplanned environmental, personal or other problems which may have interfered with the performance of the candidate
 - suspected or proven irregularities committed by the candidate
 - suspected or proven irregularities committed by any other parties to the assessment
 - any suspected or proven bribery, threats or sexual or other harassment of or by candidates
 - any grounds for doubting the authenticity of the evidence presented during the assessment process;
 - I shall also review assessment instruments in terms of validity and assessment guides or processes being followed, also paying particular attention to validity, fairness and transparency, and report on these, including collated useful comments from assessors and moderators, to the LGWSETA ETQA Manager; this feedback will be used to improve old guides and instruments and develop new ones in a continuing process of renewal and improvement of quality;
 - all relevant information concerning the quality of education and training provision will be included in my reports; this information will be included

in the team report on the provider's accreditation status and forwarded to the LGWSETA ETQA Manager;

- any constructive comments about the standards or qualifications will be gathered from assessors and moderators and collated; the results of reflections on these will be included in my reports to the LGWSETA ETQA Manager; these will be forwarded to the relevant SGB for inclusion in the review process;

I shall conduct my work in line with the vision and mission of the LGWSETA, particularly in relation to improving the quality of education, training and assessment for all learners in the sector, giving guidance and support to all learners to achieve their full potential.

Signed: _____ Date: _____

Name in full (printed): _____

Reading 5

Gipps, C., *Beyond Testing*, Falmer Press, 1994, pp. 3-4.

The first question to be asked then when considering the form of assessment to be used is 'what is the assessment for?' For example assessment to support learning, offering detailed feedback to the teacher and pupil, is necessarily different from assessment for monitoring or accountability purposes (for a start it is much more detailed). We must first ask the question 'assessment for what?' and then design the assessment programme to fit.

I take the view that the prime purpose of assessment is professional: that is assessment to support the teaching/learning process. But, government, taxpayers and parents also want to know how the education system and individual schools are performing and they must have access to that information. A major, though not the only, element of this information is pupil performance as measured by tests and examinations. Assessment carried out for these purposes is likely to be more superficial since it needs to be relatively quick and manageable and needs to be more reliable than that to support learning. One can picture it as a form of survey (using postal questionnaires) as opposed to an in-depth study (using detailed interviews). Somewhere in between these two extremes of testing to support learning or for accountability purposes lies assessment for certification purposes, as with our public exams at 16 and 18: this assessment has to be both detailed (to provide comprehensive coverage) and reasonably reliable (so that we may have confidence that the results are comparable from one school to another and from one part of the country to another) though in other countries, for example Germany, this is not seen as an issue.

The problem that we have to confront is that tests designed for purposes other than to support learning – the huge quantities of multiple choice standardized tests in the USA, and the formal written exam in the UK – have had, we now realize, unwanted and negative effects on teaching and the curriculum. The stultifying effect of public exams on the secondary system in England has been pointed out by the HMI (1979 and 1988), and was a prime mover in the shift towards GCSE with its broader range of skills assessed, a lessening of emphasis on the timed exam and an opening up of the exam to a broader section of the age cohort. (All of this was brought in and supported by the same government which is now retrenching to a formal, exclusive, written exam, but that is another story). The limiting and damaging effect of standardized multiple-choice tests in the USA has also been well documented and analysed in recent years (for example, Resnick and Resnick, 1992). But assessment for monitoring and accountability purposes will not go away; on the contrary, a number of countries in the developing world even more to gear up their education systems: in the USA, in New Zealand, in Australia as in Great Britain governments have linked economic growth with educational performance and are using assessment to help determine curriculum to impose high 'standards' of performance and, in New Zealand and Britain, countries which have taken on board the New Right marketplace model, as a market signal to aid parental choice and competition between schools (Murphy, 1990; Willis, 1992a).

Mindful of the distorting effects of assessment for these purposes, the task assessment specialists must address is how best to design accountability assessment which will provide good quality information about pupils' performance without distorting good teaching (and therefore learning) practice. We must also explore other forms of assessment which can be used alongside accountability assessment to support learning, and criteria by which we can evaluate them. This is not to say that traditional standardized tests and examinations have no role to play on assessment policy, but that we need to design assessment programmes that will do what is required of them and have a positive impact on teaching and learning.

This brings us to the second question which should be asked, but almost never is: '*what kind of learning do we wish to achieve?*' for we know now that different forms of assessment encourage, via their effect on teaching, different styles of learning. If we wish to foster higher order skills

including application of knowledge, investigation, analysing, reasoning and interpretation *for all our pupils*. Not just the elite, then we need our assessment systems to reflect that.

But a failure to articulate the relationship between learning and assessment has resulted 'in a mismatch between the high quality learning described in policy documents as desirable and the poor quality learning that seems likely to result from associated assessment procedures' (Willis, 1992b, p.1).

We need to put on the assessment agenda issues of learning style and depth. We must articulate the model of learning on which we are to base new developments in assessment over the next decade if we are to develop a sound model and one which will achieve the results we wish for it. After all, the original psychometrics was based on a theory of intelligence, while multiple choice standardized tests were based on a behaviourist model of learning: educational assessment for the next century must be based on our best current understanding of theories of learning.

Reading 6

Slippery Rock University's *Governing Principles for Assessment of Student Learning**

1. **Assessment begins with a vision of the kinds of learning and development we most value for students.**

The assessment process exists primarily as a vehicle for improving learning. The learning and development we value must govern not only *what* we choose to assess but also *how* we go about the task. Other purposes for assessment, such as accountability and public relations, are secondary and must not interfere with its primary purpose.

2. **Assessment must reflect an understanding of learning and development as multidimensional, integrated, and revealed in performance over time.**

As a complex process, learning entails not only knowledge and abilities but also values, attitudes, and habits of mind that affect both academic success and performance beyond the classroom. Assessment processes should employ a diverse array of methods, including those that call for actual performance, using them over time so as to reveal change, growth, and increasing degrees of integration.

3. **The assessment process is most effective when it is ongoing, systematic, and self-aware.**

Assessment is a process whose power is cumulative. Progress towards stated goals for learning and development should be monitored continuously, and with an eye on how a program's various elements are working together for success. The assessment process itself should be evaluated and refined in light of emerging insights.

4. **Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes.**

The assessment process entails comparing educational performance with educational purposes and expectations derived from our institution's mission, from faculty and staff intentions in program and course design, and from knowledge of students' own goals. Assessment as a process pushes us toward clarity regarding where to aim and what standards to apply; assessment also prompts attention to where and how program goals will be taught and learned. The process of assessment encourages further consideration and refinement of stated purposes.

5. **Assessment planning and implementation are collaborative activities that must involve faculty, staff, students, and administrators.**

Assessment is not a task for small groups of experts. Student learning and development is a campus-wide responsibility, and assessment is a way of enacting that responsibility. In addition, the assessment process relies on the expertise of faculty and staff concerning the learning objectives and methods endemic to their disciplines. Assessment may also involve individuals from beyond the campus whose experience can enrich the institution's sense of appropriate aims and standards for learning and development.

- 6. Effective assessment of program goals, of student learning, and of student and staff development occurs when results are disconnected from decisions concerning faculty and staff retention and promotion.**

Faculty and staff will actively and creatively participate in assessment activities only if they trust that those responsible will not use the results of such activities to justify adverse personnel decisions. At the same time, faculty and staff should know that their participation in institutional and program assessment is valued when they are considered for retention and promotion.

- 7. Assessment leads to improvement when high quality teaching, learning, and student development are visibly valued, encouraged, and implemented.**

The push to improve educational performance must be a visible and primary goal of the university's leadership, if assessment of learning is to be effective. Improving the quality of undergraduate education must be central to the institution's planning, budgeting, and personnel decisions. Resources must adequately support professional development regarding assessment, as well as the creation and implementation of assessment plans. In addition, the innovations, alterations, and activities undertaken by departments and programs as a result of the assessment process must be seriously considered when resources are allocated.

** While some of this draft of Slippery Rock's assessment principles is original, much has been borrowed from the "Principles of Good Practice for Assessing Student Learning," composed by several experts in the field, under the auspices of the AAHE Assessment Forum and with support from the Fund for the Improvement of Postsecondary Education.*

Reading 7

Kramer, D., *OBE Teaching Toolbox*, pp. 58-76.

PERFORMANCE ASSESSMENT TECHNIQUES

There are a number of basic approaches to assessing performance or performance tasks. In this section we will have a closer look at each one. It is important to remember that there are different ways to assess a task and creative educators can design even more. The techniques discussed below should not be seen as a complete list.

RUBRICS

One of the problems of giving a quantitative mark, such as a percentage, in assessment is that such a mark gives not information to the learner about what, in particular, was good or bad about their work. We all know that 80% is a good mark for a piece of work but we cannot say exactly what 80% is. Worse still, the same piece of work, marked by 10 different educators, could conceivably be given 10 different scores, even if they all give almost the same marks. It would be far better to have assessment instruments that give clarity about the quality of performance.

A rubric is an assessment tool that describes a continuum of performance quality that ranges from poor to excellent. A rubric lists each of the elements needed to perform a task and then describes exactly what constitutes acceptable performance for each element. The rubric also describes excellent performance that exceeds the level of acceptable work and inadequate performance that does not yet meet the minimum level of performance. Usually rubrics use a grid or table format, but there are many different ways of creating a rubric. As with all instruments, we can adapt them to our individual styles and needs.

Rubrics are therefore analytical rather than quantitative. They analyse the performance rather than giving it a mark. Clearly, rubrics are valuable in giving learners information about their work and are therefore useful for self and peer assessment. It is important to understand that work which falls short of the expected or required level of achievement is not work for which the learner gets assessed as having failed. It means that the learner has to try again to improve those elements which are not yet acceptable. Where this is not or cannot be done, for whatever reason, the learner cannot be credited as having achieved the standards.

Another important point is that a rubric sets out the required standard for each of the elements of the task. What happens if the learner achieves the standards in the case of 90% of the elements but is below standard in 10% of the elements? In other words, the task is almost properly done but not completely. In traditional assessment educators would be happy to say that the learner has mastered the task at a level of 90% - a very good assessment. Let's ask the question differently. How much unacceptable work is acceptable? In performance assessment, the required standard must be achieved in each of the task. The rubric helps us to underline this principle.

HOW TO DESIGN A RUBRIC

There are many ways to design a rubric, some of which are very technically demanding. We will look at the basic steps.

Decide what the task is and what kind of evidence of performance should be assessed. Decide on the elements of the task. What are the different parts of the performance that will be

assessed? What knowledge must be mastered? What skills must be used or actions be taken? About what issues, should personal opinions, values or insight be expressed? What will make the task successful in respect of each of these issues? The educator alone could decide this but it would be better if the educator worked with the learners to agree on the elements that will make a successful task.

1. Condense the most important ideas into a shorter list of main parts of the task. Assessment of a task must be able to be carried out by learners or the educator. We do not want a complex and cumbersome instrument for assessment and so we need to ...
2. Identify the most significant performance criteria that we will look at in assessing the task.
3. Define each of the performance criteria. Each criterion needs certain actions to be taken by the learner in order to achieve success. What does each criterion mean in practice? We need to state the performance criterion as an action that can be assessed.
4. This means describing the different levels of performance, excellent, acceptable and poor. We need to expand on what we would consider to exemplify the different levels of performance in each category. First we need to describe the benchmark – that level of performance that is the minimum for the performance to be acceptable. After this we can describe excellent performance and non-acceptable performance.
5. Add the detail for each level of performance. We fill in all the details so that anyone can understand.
6. Lastly, we need to check and verify the final rubric. This means applying and testing the criteria to see if they work. We need to check with other people and against the original intentions of the task.

EXAMPLE 1: RUBRIC FOR A MATHS PROJECT

	ACCEPTABLE WORK		UNACCEPTABLE WORK	
	EXCELLENT	GOOD	NEEDS WORK	POOR
MATHS KNOWLEDGE Using maths to solve problems. Applying maths concepts, principles & processes	Shows complete understanding of concepts/principles. Uses proper terminology/notation. Applies algorithms and processes correctly.	Shows nearly complete understanding of concepts & principles. Uses nearly correct terminology/ notation. Algorithms & processes generally correct but with minor errors.	Shows limited understanding of concepts or principles. Misuses or fails to use correct terminology or notation. Major maths errors.	No understanding or no answer attempted.
STRATEGY & PLANNING Identifying the important elements of the problem. Using models, diagrams & symbols to show the relevant concepts.	Identifies all important elements & shows complete understanding of their relationships. Creates an effective strategy to solve problem. Solves the problem effectively.	Identifies most of the important elements of the problem & shows general understanding of their relationships. Creates an acceptable strategy to solve the problem. Solves the problem acceptably.	Identifies some of the important elements of the problem & shows limited understanding of their relationships. Shows limited or unclear strategy. Gives some evidence of solution.	No useful strategy.
COMMUNICATION SKILLS Presentation of the explanation and rationale for the solution to the problem.	Gives a complete and clear explanation of the solution process showing what, how and why it was done. Uses diagrams or graphics where appropriate.	Gives a nearly complete explanation of the solution process but with minor gaps. Sometimes uses diagrams or graphics, where appropriate.	Gives inadequate or unclear explanation of the solution process. Uses some diagrams or graphics but in an inappropriate or unclear way.	No or unclear explanation.

EXAMPLE 2: RUBRIC FOR SCIENTIFIC INVESTIGATION

	EXCELLENT	ACCEPTABLE	NOT YET ACCEPTABLE
USE OF DATA	All appropriate data is correctly collected, organised and analysed with accurate and relevant interpretations.	Most of the relevant data is collected. Data is generally organised and acceptably analysed. Some useful interpretation.	Important data is missing. Data is poorly organised or confused. Data is not analysed or analysed inadequately.
SCIENTIFIC KNOWLEDGE	Shows mastery of and explains all scientific concepts and principles. Uses correct symbols and terms.	Shows mastery of and explains some scientific concepts and principles. Uses mostly correct symbols and terms but with minor errors.	Shows poor mastery of scientific concepts and principles. Uses incorrect symbols and terms or uses them incorrectly.
SCIENTIFIC PROCEDURES	All necessary processes and procedures safely, precisely and effectively followed. Equipment properly chosen and handled.	Most necessary processes and procedures generally safely and effectively followed, but with some flaws. Equipment chosen and used adequately.	Necessary processes and procedures unsafely or inadequately followed or not used at all. Equipment incorrectly chosen and used.
PRESENTATION	Observations well described. Results are effectively interpreted and logically explained. Effective use of clear graphics, diagrams and statistics.	Observations adequately described. Results are adequately interpreted and explained. Some use of graphics, diagrams and statistics.	Observations poorly described. Results are inadequately interpreted and explained or confused and illogical. Poor use of graphics, diagrams and statistics.
GROUP WORK	Shows a high degree of collaboration, sharing of workload and knowledge. Supports and helps group members very much.	Shows some degree of collaboration, sharing of workload and knowledge. Gives some support and help to group members.	Shows no collaboration, sharing of workload or knowledge. Gives no visible support or help to group members.

The above examples can be adapted for most kinds of task, including both natural, physical and social sciences. These two examples use a table to show the different levels of performance. Another layout for a rubric uses a simple list, as shown below:

EXAMPLE 3: RUBRIC FOR A WRITING ASSIGNMENT

The Task: The School Governing Body is soon to decide whether to do away with the wearing of school uniforms. Write to the principal of the school to put your point of view and to convince him/her of your argument.

Excellent: Well structure response with fully developed reasons elaborated in great detail. Strong organisation with much evidence of planning. Strong introduction and conclusion with ideas logically and clearly laid out. Technically competent with no errors in spelling, grammar or use of vocabulary.

Good: Structured response with reasons elaborated in detail. Good organisation with some evidence of planning. Good introduction and conclusion with ideas logically and clearly laid out. Technically competent with up to 6 errors of spelling, grammar or use of vocabulary.

Acceptable: Structured response with elaborated reasons. Organised with some evidence of planning. Acceptable introduction and conclusion. Logical and clearly laid out ideas. Technically acceptable but with more than 6 errors of spelling, grammar and use of vocabulary.

Unacceptable: Unstructured. No explanation of reasons. Inadequate organisation. Inadequate evidence of planning. Ideas illogical and not properly laid out. Confusing. Technically inadequate with more than 10 errors of spelling, grammar or use of vocabulary.

PERFORMANCE TASK LISTS

The performance task list is a list that sets out all the actions that the learner needs to attend to in completing the task. Each of the important elements of the task is listed and this allows the learners to see at a glance what they need to do. The learner is then able to work independently in doing the task and also to do self assessment by referring to the list. Another advantage is that parents are also able to see what their children are expected to do and so they, too, can be made aware of and can participate in the learning process and assessment.

The list helps educators and learners to focus on the specific elements of learning, one by one. Educators can accurately and consistently evaluate learners' work and point out specific areas of strength and weakness.

A third benefit of the list is that it allows educators and learners to weight or prioritise the different activities or elements of the task. The most important elements of the task carry a heavier weighting and will contribute more towards the final score. The elements of the task that are less critical carry a lighter weighting and will contribute less towards a final score. We can now ensure that learners understand which parts of the task are most important and should enjoy most attention.

Fourthly, a strong benefit of the task is that educators and learners can create a list by jointly analysing a task and agreeing what needs to be done. The discussions can include how the different parts of the task should be weighted.

There are many variations in the way that task lists are constructed. There are simple lists for younger learners that have smiley faces to indicate how the learner has done. There are sophisticated lists for older learners that give detailed instructions and descriptions of each part of the task, with weightings and other constraints.

Many educators prefer to use task lists rather than rubrics as lists allow a score to be given.

HOW TO CREATE A TASK LIST

We have already looked at the various elements that go into creating a performance task. These elements need to be combined into a description that tells learners about the task. We can do this by giving learners firstly, a briefing sheet that sets out the details of the task and, secondly, a task list which lists each action that the learner must take.

The **BRIEFING SHEET** sets out the following information:

1. **Background for the task.** In this paragraph we set out the context, setting or situation for the task. Where possible this should be an authentic or real situation or role that the learners can identify.
2. **Task.** This is a clear, short statement of what the learner needs to do. (You must.....)
3. **Purpose.** This shows what the task intends to achieve or why the task is set. (So that.....)
4. **The audience.** This shows whom the task and purpose are aimed at. It provides information about how the task should be presented, the level of sophistication and other clues about what would be acceptable in a real context.

5. **The procedure.** This explains the steps that must be followed in doing the task.
6. **Materials and equipment.** We may need to list what the learner will need to use to accomplish the task. We may want to restrict the learner to certain items.
7. **Deadlines.** This tells the learner when the task must be completed and ready for assessment. We may set a series of times or dates by which work-in-progress must be shown.

EXAMPLE1: TASK LIST FOR A DESIGN PROJECT

BRIEFING SHEET

BACKGROUND

The school needs to build extra classrooms to accommodate the growth in the number of learners in the area. The Governing Body estimate that they have enough money to build 3 new classrooms which are the same size as existing classrooms.

TASK

Create a design for a classroom such that 3 of these can be built on the existing school properly.

PURPOSE

The purpose of the task is to show that you can design a plan according to the measurements and constraints laid down by the Governing Body.

AUDIENCE

The governing body of the school.

PROCEDURE

1. Measure the dimensions of existing classrooms.
2. Draw a plan showing all dimensions.
3. Draw a map of the school with scale, legend, dimensions and orientations. All dimensions should be in metres.
4. Design a rectangular classroom plan, giving all dimensions.
5. Make a rough draft of the map with 3 new classrooms placed on it.
6. Ensure that there is at least 10 metres between all classrooms for safe access.

MATERIALS

Use paper and pencil to do all drawings.

The TASK LIST is the actual assessment instrument that lays out, in separate columns, the following details.

1. **The elements of the task that will be assessed.** Each assessable element is stated in a way that allows a score to be given.
2. **The maximum possible points** for each of the elements.
3. **The weighting of each element.** Where weighting is used, and this is only an option, the most important elements carry a heavier weight than the least important ones. This allows the learner to see which are the most critical elements of the task. An option is not to use weighting at all, but to simply take 100% as the total maximum possible score and then to divide the maximum possible scores between the elements so that the most important elements have a higher maximum possible score.

4. **The score awarded by the learner** as part of self assessment, where this is part of the process.
5. **The actual score awarded by the educator** in the assessment of the task.

EXAMPLE 1: TASK LIST FOR DESIGN PROJECT			
TASK LIST			
ELEMENT	POINTS	SELF	EDUCATOR
1. All measurements taken accurately	10		
2. An accurate classroom drawing is made	10		
3. All dimensions are accurately drawn to scale	15		
4. A scale map of the school is drawn	20		
5. A legend, scale and orientation are included	10		
6. All major features are included on the map	15		
7. A draft map with 3 new classrooms is drawn	5		
8. There is at least 10 metres between classrooms	5		
9. All features are well labeled	5		
10. All work is neat and well presented	5		
TOTAL:	100		

In the above example each element has a score that shows its relative importance as part of the project. A different approach is shown in example 2. Here, all the elements have a range of possible scores, from 1 to 4.

EXAMPLE 2: TASK LIST FOR BIOLOGY STUDY	
BRIEFING SHEET	
BACKGROUND	
You have been asked by an educator to help teach a Grade 6 class about how animals are adapted to live in specific environments (eg. deserts, forests or mountains). The class has just been learning about different parts of the world and their geographic features.	
TASK	
Write a research report about how three different animals are adapted for living in their environment.	
PURPOSE	
The purpose of the work is to provide the educator with the information needed to teach Grade 6 learners about biological adaptation of animals in different environments.	
AUDIENCE	
Grade 6 learners, many of whom are second language speakers.	
PROCEDURE	
<ol style="list-style-type: none"> 1. Select 3 animals, each of which lives in a different environment. 2. Research and discuss how each is adapted to its environment. 3. Use pictures or drawings to illustrate your report. 	

EXAMPLE 2: TASK LIST FOR BIOLOGY STUDY				
TASK LIST				
	ELEMENT	POINTS	WEIGHT	SCORE
1	The concept of biological adaptation is explained.	1 2 3 4	1 2 3	
2	The relationship between structure and function is explained.	1 2 3 4	1 2 3	
3	Appropriate biological terms are used and explained.	1 2 3 4	1 2 3	
4	At least 5 features per animal are discussed.	1 2 3 4	1 2 3	
5	The report is well laid out in logical paragraphs.	1 2 3 4	1 2 3	
6	The report is neat, well presented and attractive.	1 2 3 4	1 2 3	
7	Pictures and diagrams are relevant and useful to the explanations.	1 2 3 4	1 2 3	
8	Diagrams are labeled and captioned.	1 2 3 4	1 2 3	
9	All the major adaptations of each animal are accurately explained.	1 2 3 4	1 2 3	
10	The language and presentation are appropriate for Grade 6.	1 2 3 4	1 2 3	
	TOTAL:			

In this case the elements are weighted either 1, 2 or 3. The educator circles the points for each element and multiplies this by the weighting to get a total score.

Those with heavier weightings will therefore carry a higher total score. It is important to ensure that learners know the weighting before they begin the task, so that they know which elements are the most important for the successful completion of the work.

RATING SCALES

A rating scale is another variation of a rubric and a task list in that it sets out the performance activities and products together with a clear statement of performance standards. The rating scale uses points on a continuum to score the quality of the work. Conventionally, low scores indicate poorer quality and high scores indicate better quality. The key to the rating scale is that each score is defined and we can refer to a key to see exactly what that score means. The benefit of scales is that they allow a number of related but separate performance indicators to be grouped and assessed together. We can assess skills, knowledge and attitudes together or in parallel or separately. Another benefit is that, because rating scales record a shorthand symbol such as a number or statement, or for a detailed list of knowledge, skills or values items, they help with simple but accurate recording in large classes. Scales may use numbers (1,2,3....) or letters (A,B,C.....) or words (Excellent, Good, Average, Poor, Very Poor) as symbols. Some scales use icons or pictures, such as smiley faces, to show the level of achievement. This is particularly useful when dealing with very young learners whose reading skills may not yet be well developed.

The key to a good scale is to keep it simple. The statements should be short but clear enough for learners to use in self or peer assessment. Compiling a rating scale requires three actions:

- deciding and grouping the things that need to be assessed.
- deciding on how many levels of performance there will be and their symbols.
- defining the characteristics for each level of performance.

There are three basic types of rating scale:

ACHIEVEMENT SCALES

These are similar to the example given above and are used to describe only one element of performance. This means that there would need to be as many achievement scales as there are performance indicators. Below are various examples of achievement scales. The scale can be used by assigning a score for each of the indicators.

EXAMPLE 1: RATING SCALE FOR MATHEMATICS

	PROBLEM SOLVING	WORK PRESENTATION
3	Understands problem easily. Shows plans and uses appropriate problem solving techniques.	Logical, clear, and accurate work. Neat, appropriate, detailed, precise drawings and graphs.
2	Understands the problem. Uses appropriate problem solving techniques.	Mostly logical, clear and accurate work. Neat, appropriate drawings and graphs. Can be more precise.
1	Understands problem in part. Plans and uses problem solving techniques with help.	Logical, clear and accurate work in part. Drawings and graphs acceptable but can be improved.
0	No understanding of problem. Unable to plan or use problem solving technique.	Illogical, unclear or inaccurate work. Untidy or inaccurate drawings and graphs.

EXAMPLE 2: RATING SCALE FOR SCIENCE INVESTIGATION

Equipment:

Chooses and uses the right materials and equipment safely.

SCORE
0 1 2 3 4

- 4. Selects correct requirements and uses properly & safely.
- 3. Selects mostly correct requirements and uses properly.
- 2. Selects some requirements and needs help in handling them.
- 1. Needs help in selecting requirements and in using them.
- 0. Has no understanding of the requirements and handles them badly.

Observation

Observes carefully and records observations and data accurately.

- 4. Observes methodically and records relevant data systematically and neatly.
- 3. Observes carefully and records relevant data in an orderly fashion.
- 2. Observes and records most relevant data.
- 1. Makes cursory observations and records only some relevant data.
- 0. Observes poorly and records haphazardly, untidily and/or illogically.

0 1 2 3 4

Analysis

- 4. Faultless procedure, intelligent analysis. Conclusion logically based on data.
- 3. Good procedures and analysis of data. Conclusions drawn from data.
- 2. Acceptable procedures, analysis and conclusions. Needed some help.
- 1. Achieved acceptable conclusions after much help in analysis.

0 1 2 3 4

There are many variations in both layout and the way that we can describe levels. One of the more interesting examples categorises the levels as ranging from Novice (i.e. low achievement), to Intermediate, Advanced and lastly, Superior.

EXAMPLE 3: ORAL READING RATING SCALE

- Level 4: Read in larger meaningful phrases. Some deviation from text, repetition, hesitation or mispronunciation, but not enough to disturb the presentation. Reads with expression and animation. Skillful use of intonation, pitch and timbre.
- Level 3: Reads in phrase groups of up to 4 words. Noticeable deviation, hesitation and pronunciation errors although the sense of the writings is preserved. Reads with expression but still somewhat mechanically. No personal interpretation.
- Level 2: Reads mainly in 2 word phrases. Significant lack of fluency and little adherence to syntax. Little expression and presentation unrelated to content.
- Level 1: Reads mainly word by word. Reads with hesitancy, frequent pronunciation errors and little overall preservation of structure or syntax. Poor word recognition.

FREQUENCY SCALES

These assign scores to the performance based on how often the learner performs at each level of quality. The intention is to give a picture of which level most characterises the work. These scales are useful where assessment is applied to learning activities that span a period of time or where we are assessing a number of examples of the learners' work.

EXAMPLE 4: RATING SCALE FOR SPEAKING SKILLS AND DEBATING

	1	2	3	4	5
Shows evidence of planning and research on the topic.					
Prepares effective visual aids to support presentation.					
Delivers effective introduction and conclusion.					
Delivers argument without hesitation or interruption.					
Has strong, relevant, interesting and compelling content.					

1: never 2: rarely 3: sometimes 4: often 5: always

The above example is a frequency scale for assessing debating or speaking skills. Similar scales can be developed for outcomes such as teamwork, leadership, time management and many other outcomes that are mainly psychomotor or affective in nature. These scales are particularly useful to assess observable values and attitudes.

HOLISTIC RATING SCALES:

Perhaps the most useful type of rating scale but also the most complicated is the holistic scale. As its name implies, the holistic scale combines many elements of performance. Each level represents an overall picture of the quality of the performance, combining all the indicators into one score. The assessment assigns a score that most accurately describes all the performance indicators at once, rather than focusing on each indicator, one at a time.

Holistic scales are most useful when assessing work that requires an extended range of skills, knowledge and attitude items, over an extended period of time. They are also useful to assess tasks that require the learners to combine selected skills and to recall knowledge from a wide

range of learning. The biggest advantage of holistic scales, however, is that they can be used to assess integrated work. If, for example, a programme integrates language, life orientation and technology, a holistic scale can be used to give an overall assessment of the performance at once. This makes the holistic scale more useful in assessing final demonstrations, rather than small tasks undertaken during the learning process.

These scales take a lot of time, thought and effort to develop but are exceptionally useful. A second disadvantage is that learners rarely perform at the same level as regards all the performance indicators and so the scales tend to “average out” the performance indicators.

The example below is of a holistic scale for assessing an essay. In this example an educator could use either a numeric score (1,2,3, etc.....) or an alphabetical score (A,B,C etc.....).

Holistic scales can be used for all learning areas and can be as detailed as needed. A feature of the scales is that they do not refer to specific exercises or problems but rather to the overall performance of the learner. Example 7 below can be adapted for science, humanities, maths or other areas of learning quite easily because it covers broad issues.

EXAMPLE 5: HOLISTIC RATING SCALE FOR AN ESSAY

1	2	3	4	5	6
C		B		A	A+
C+		B+			
Numerous	Deficiencies	Generally	Acceptable	Generally	Excellent

6. (A+)

- Well organised, creative, original and stylish work.
- A strong and articulate personal voice, with the learner’s real thoughts and feelings showing clearly.
- Well planned and organised work. Effective structure, with a clear compelling introduction and a conclusion that is both consolidating and personal in opinion. Competent sentence construction. Vocabulary is well chosen, vivid, interesting and appropriate. There are no spelling or grammar errors.
- The content makes interesting reading. The topic has been well researched and the work has plenty of illuminating details to support the central idea. The work is enjoyable to read.
- There is a logical, well considered argument that shows a clear grasp of the issues, problems and solutions.

5. (A)

- Well organised and creative work
- A strong personal voice, with the learner’s real thoughts and feelings often evident.
- Good planning and organisation. Effective structure, with an interesting introduction and a conclusion that summarises the ideas. Competent sentence construction. Appropriate vocabulary that is interesting and appropriate. There are few spelling or grammar errors.
- The content makes interesting reading. There is evidence of research and sufficient detail to support the central idea. The work is enjoyable to read.
- There is a logical, considered argument that shows a clear grasp of the issues, problems and solutions.

4. (B+)

- Organised work with examples of creativity.
- Personal voice often shows the learner’s real thoughts and feelings.
- Planning is evident and the work shows some organisation. Satisfactory structure, with an introduction and a conclusion. Competent sentence construction. Vocabulary is interesting and appropriate. There

are some spelling or grammar errors and the work has some detail to support the central idea. The work is pleasant to read.

- There is a logical argument that shows some thought and a grasp of the issues, problems and solutions.

3. (B)

- Work shows some organisation but deals with the topic at a superficial level.
- Little personal voice which occasionally shows the learner's real thoughts and feelings.
- Some planning is evident but the work lacks organisation. Structure is adequate. There is an introduction and conclusion. Immature sentence construction. Vocabulary is sometimes inappropriate, immature or vague. There are many spelling or grammar errors.
- The content is relevant and sometimes interesting. There is some evidence of research but lacks sufficient detail to support the central idea.
- The argument shows some logic. The learner shows a shallow grasp of the issues, problems and solutions.

2. (C+)

- Work shows little organisation. Learner attempts to focus on the topic.
- No personal voice.
- No planning is evident. Structure is weak and lacks a distinct introduction and conclusion.
- Sentence construction shows flaws. Vocabulary is inappropriate, immature or vague. There are many spelling or grammar errors that make the work difficult to understand.
- The content is irrelevant and often off the topic. There is no evidence of research or any detail to support the central idea.
- The argument shows no logic. The learner shows a little grasp of the issues, problems and solutions.

1. (C)

- Work shows no organisation. Learner has not tried to focus on the topic.
- No personal voice.
- No planning is evident.
- The work is unstructured and lacks any introduction or conclusion. Sentences show no mastery of elementary structuring. Vocabulary is inappropriate, elemental and often incorrect. There are many spelling or grammar errors that make the work difficult to understand.
- The content does not relate to the topic. There is no evidence of research.
- The argument is illogical. The learner shows no grasp of the issues, problems and solutions.

EXAMPLE 7: HOLISTIC RATING SCALE FOR A RESEARCH PROJECT

- A: The project (or research) is of an excellent standard. The work is correct, relevant, complete and appropriate in every respect. Quality planning and intensive research are evident. There are clear and useful illustrations, statistics and graphics to support the work. There is evidence of higher order thinking and of a personal approach to the design of the work. Minor errors do not affect the substance of the work.
- B: The project (or research) is of a high standard. The work is generally correct, relevant, complete and appropriate. There is evidence of planning and significant research. The work includes useful illustrations, statistics and graphics. Some higher order thinking and personal interpretation are evident. There are, however, some omissions and inaccuracies that affect the overall quality of the work.
- C: The project (or research) is of an acceptable standard. The work is relevant and generally correct but it contains omissions, inaccuracies as well as incomplete, incorrect or inappropriate material. The presentation lacks appropriate supporting illustrations, statistics or graphics. There is some evidence of misconception about the topic.
- D: The project (or research) is unsatisfactory and falls short of an acceptable standard. The topic is not addressed sufficiently or correctly. There is no evidence of planning or research into the issues.

The presentation is unordered and lack appropriate supporting graphics or statistics. There are serious misconceptions and no evidence of higher order thinking.

A simpler version of a holistic rating scale requires the educator to mark the learner's achievement along a line, as in example 8. Here the educator simply places a mark on the line in a way that best shows how the learner has done.

EXAMPLE 8: RATING SCALE FOR AN ESSAY

PERFORMANCE

RATING

Does the writer give a clear introduction which states the main idea?

Clear Somewhat Clear Needs Improvement

Does the writer give a clear, summarising conclusion?

Clear Somewhat Clear Needs Improvement

Does the writer use sentences, reasons or arguments to explain, support or develop the main idea?

Supportive & Organised Some Support Needs Improvement

Does the writer use correct grammar and spelling?

Correct Some Errors Needs Improvement

Does the writer use expressive, well considered, appropriate language?

Well Considered Somewhat Considered Needs Improvement

Reading 8

Pahad, M., *Approaches to Assessment and Evaluation*, Wits/IEB coursebook, pp 34-37

Norm Referencing or Criterion Referencing?

With norm-referenced tests, the intent is to compare student performance on the test with that of a norm group of students, rather than to determine how proficient a student is in a particular subject or skill.

(Capper, 1993)

Norm referencing places candidates in rank order after assessing them (typically through an external examination) at a particular point in time (eg, after ten years of school), and awards distinctions to a certain percentage at the top, and other grades in descending rank order, until a certain percentage at the bottom are left with failure. Sometimes the pass rate or distinction rate is fixed as a predetermined percentage. For example, many countries select students at the end of primary school for a given number of places in the more academic schools in a two-tier secondary school system, or a limited number of secondary school places. Perhaps the 'top' 10% pass and the rest are doomed to failure. If standards were rapidly improving, as is sometimes the case in developing countries, an individual could fail even though he or she had done better than someone who had passed the previous year.

Criterion referenced tests record positive achievements at different levels. We are all familiar with certificates awarded for playing musical instruments, or for different levels of competence in ballet or swimming. If you reach the required standard, you pass. If you have not yet reached it you must try again. Criterion-referenced school examinations try to provide measurable criteria so that a student's progress can be monitored and his achievements acknowledged in a similar manner. Anyone who meets these criteria is credited with the appropriate level pass in that subject. If 100% of candidates were to meet the criteria they would all pass.

So the shift towards criterion referencing reflects a desire to move away from an assessment system which is primarily designed to select towards one which is primarily designed to credit achievement at different levels. Criterion referencing also makes the assessment criteria quite explicit, so that every student should be able to understand how her work is to be assessed. Similarly, every teacher should be able to explain results with reference to stated criteria. Clearly criterion referencing is more philosophically in tune with a society which aims to give the whole population a good general education and stresses the ideas of access, equity and redress, as does ours.

Think about the adjustment of marks.

- With norm referencing the statistical adjustment of marks causes enormous unease and suspicion; and indeed some methods can be crude, like forcibly imposing a 'normal curve' on unruly graphs reflecting the raw scores, regardless of unusual circumstances. Some adjustments, however, are both logical and fair. Suppose that the raw scores from an examination set for years by the same board for students with approximately the same profile from a stable group of students showed that half the students had failed and no single candidate had gained a distinction. Supposing that in that subject no more than 10% had failed in any previous year and between 7 and 9% had always achieved distinctions, what would you think? Could it not be explained by the fact that the questions were too hard, or were not set on the agreed syllabus, or that the marking memo had been seriously at fault?
- Take another example: a teacher sets an end of year examination for a class of children all of whom she expects, from their successful year's work, to be promoted and to do well in the class above. To her surprise they all score very much lower than expected and

several fail. Would the teacher accept the raw score and recommend that the children repeat the year?

Reasonable educators in such circumstances would investigate the cause of the deviation from expected results, and they would also be bound, out of fairness to the candidates, to adjust the raw scores, whilst maintaining their rank order. **Raw scores merely give the tester information which still needs interpretation.** Such scores are not sacrosanct and should not be decontextualised. Marks may need to be adjusted even in criterion referenced tests when it is perceived that an unintended factor has unfairly influenced the results.

But how do we move towards criterion referencing? We must define a number of learning outcomes at various levels and then set about recording each student's progress in relation to criteria which demonstrate competence or mastery at the various levels.

Defining the learning outcomes and breaking them down into specific areas of content to be understood and skills to be developed has been problematic for all countries moving in this direction. The tendency is to move away from the British model which has been criticised for **fragmenting learning into thousands of tiny discrete attainment targets, which lost meaning as they lost context.** When we try to isolate items of knowledge and understanding which are comparatively simple to test, we often merely succeed in isolating items whose significance is minimal in terms of real progress in learning. Broader bands of related skills/knowledge which represent aspects of the development of complex higher order cognitive skills/concepts/understanding are now preferred.

The complex higher order cognitive skills which must be developed demand the ability to select and combine many different skills and concepts, and much diverse information, and to transfer their application to unfamiliar situations and new problems. Thus thematic approaches with realistic and relevant contexts which integrate the curriculum are more suitable than artificial compartmentalisation. In terms of assessment criteria this has led increasingly to the use of tables describing in blocks the kinds of things a student **should be able to do or demonstrate understanding of** at a number of levels. These are being constantly revised and each education department/examining board/school district seems to find it necessary to invent its own. If this is seized as an opportunity to empower and involve teachers, it may prove to be the best approach.

Is criterion referencing then diametrically opposed to norm referencing? Are they mutually exclusive? In reality, the distinction is not so simple, as Keith Lewin points out:

The simplest definitions of norm and criterion referenced tests distinguish between tests which are designed to compare the performance of individual students with that of other students, and those which assess performance against a criterion independent of the performance of others. A moment's thought indicates that things are not so simple.

This common kind of definition does not in fact identify two different types of tests very effectively – criterion referenced test items may look very similar to norm-referenced items. If there is a useful distinction it is more concerned with the differences in the interpretation of performance on test items than on the nature of the items themselves.

We should note that tests constructed to produce norm-referenced outcomes can be regarded as criterion-referenced – the selection of particular items creates the criteria. Similarly criterion referenced instruments can be used to rank candidates and compare their performance and may therefore be treated as norm-referenced instruments. In either case it can be argued that the validity, reliability and suitability of a test may suffer if it is designed for one purpose and used for another. However the point stands that it is the use made of the results rather than the nature of the items that most clearly distinguishes between test types.

(Lewin 1994)

Whilst the point made above is a fair one, it is also important to note that criterion referenced tests are always set and marked against known explicit criteria, whereas, although norm referenced tests **may choose** to utilise this approach, they might also choose **not to make the criteria explicit** and to shroud the process in secrecy and mystification. In addition, Lewin is discussing standardised test items; our more holistic and varied approach to summative evaluation gives a lot more scope to use more fully the opportunities provided by criterion referencing.

In conclusion, it seems clear that the primary purposes of norm referencing and criterion referencing differ greatly, and the change towards the latter is a genuine and significant step forward.

Reading 9

PEER ASSESMENT IN GROUP WORK

By Teresa Bulman
tb@caesar.geog.pdx.edu

Six years ago when I began asking students to do group projects, they balked. They still do. Their primary concerns seem to be two-fold: (1) given their busy lives, filled with job and family responsibilities, it is difficult for them to find time to meet in groups outside of class, and (2) students are leery of having their individual grades dependent upon the abilities and efforts of others.

I have addressed the first concern by allocating a small portion of class time for group meetings. The second concern has been more difficult to resolve. After several false starts, which ranged from no peer evaluation to oral peer evaluations, I perused articles in the journal *Assessment and Evaluation in Higher Education* that discussed written peer assessments. I modified some of the materials I found to create a framework for peer assessment and have used it extensively in classes ranging in size from 15 to 55 students (Sample Peer Assessment Framework). I make slight adjustments to the form each time I use it in order to fit the particular student project, but the basic scope and format remain the same. It is designed to give students (1) the opportunity to assess the contribution of each other member of the group; (2) practice in assessment; and (3) both qualitative and quantitative assessment formats so they can see the difference between the two and so they can have thinking and writing exercise (because the quantitative assessment must be supported by qualitative statements).

I have found that the peer assessment forms work as both peer pressure and as a release valve. Students know going into the project that they will evaluate their peers and be evaluated by them. This causes them (they tell me) to work harder on the project than they might if their grade only were at stake. On the rare occasion when a student does not do a fair share, the other members of the group have an opportunity to reveal that problem.

There are a few aspects of group work that I have found critical to success in the groups and in the peer assessment, and these aspects must be brought to the student's attention at the beginning of the projects.

Students need to be made aware (1) of the importance of group projects as a part of their intellectual development (researching, critical thinking, and writing skills) and as part of workforce training; (2) that goals and objectives need to be identified for any project and are particularly important in focusing group activities and assignments; (3) that organizational skills are important (assignment of individual tasks and implementation of time-tables); and (4) that every individual has a responsibility to note and address problems within the group.

Students still balk at group work, but not as vociferously as before. I'm working on doing a better job in explaining the importance of group work, and I am helped in this by employers who increasingly list ability to work in groups or teams as a job criterion. The peer assessment has provided students the opportunity to have their individual contributions and talents noted within the context of a group project.