GTIP Orientation Piece – Assessment (Graham Butt, 2008)

Graham Butt (School of Education, University of Birmingham) helps you explore assessment with trainee teachers, in terms of the background - what we understand by assessment, recent developments in assessment, assessment planning, and the fundamental question of making judgements using criteria and/or norm referencing.

Menu
- Introduction, p1
- What is assessment for? P2
- Assessment in practice, p2
- Key policy developments, p2
- Planning, p4
- ‘High stakes’ and ‘low stakes’, p4
- Norm- and criterion-referenced assessment, p4
- Conclusion, p5
- Terminology, p5
- Bibliography, p6
- Journal abstracts, p7

Introduction
The standards for initial teacher training (TTA/DFES, 2003) place considerable emphasis on trainees developing an understanding of assessment - a term which covers monitoring, assessment, recording, reporting, and accountability (Harlen et al., 1992).

Assessment is also prominent in teachers’ lives. On the one hand they feel judged by results (e.g. GCSE A-C grades), and yet on the other are encouraged to employ the techniques of Assessment for Learning (AfL).

What is assessment for?
Possible answers include:
- to enhance students’ learning
- to measure (or possibly raise) standards
- to check teaching objectives against learning outcomes
- to recognise and plan for students’ learning needs
- to place students against different descriptors of achievement
- to discover what students know, understand and can do
- to help plan future learning objectives
- to help students to devise personal targets
- to evaluate teacher effectiveness and performance
- to motivate teachers and students.
Some of these purposes are in conflict. It may be worth asking trainees to identify tensions in order to show that assessment is never a neutral activity. It is best to be clear what primary purpose we hope to achieve in practice.

**Assessment in practice**
Assessment in practice is usually based on the ‘measurement’ of an individual’s learning at a particular point in time, or on a particular task. The performance is assumed to provide a good indicator of that person’s knowledge, understanding and skills.

Sometimes this **summative assessment** is enough. However, assessments can also be used to signpost future learning goals – **formative assessment**. There is a major difference between assessment of learning (summative) and assessment for learning (formative).

Thus, assessment should not only be about end of course testing, in which students’ attainment is acknowledged by a single mark or grade. The process should also be ‘two way’, with students regularly included in discussions about their future learning. This has the advantage that teachers ‘get to know’ students and enables students more fully to understand the process of teaching, learning, assessment and improving attainment.

It is worth emphasising that the process of ‘getting to know’ students starts with learning names. There is no way to avoid the need to learn names, which may mean adopting strategies to help. For example, seating plans, or memory tricks like learning by association.

**Key policy developments**
**The Education Reform Act**
ERA 1988 represented a watershed in terms of assessment practice in England and Wales. The national curriculum provided a detailed specification of what should be taught, but from the early 1990s it became apparent that national testing of the curriculum would be both unwieldy and prohibitively expensive.

**Level descriptions**
Geography teachers were charged with the responsibility of providing a ‘level’ for each student’s attainment at the end of key stage 3. These have determined as a ‘best fit’ teacher assessment against published level descriptions. The misuse of level descriptions spread: they became pivotal to the planning and evaluation of lessons, the construction of learning activities and marking homework. They are best used only as a guide to the professional judgement of student attainment at the end of the key stage. However, the consistency of the assessment standards applied by teachers was, and is still, an issue.
Suggested activity

- Fairly late in the year (when some experience can be drawn upon) is to have your group sit in small groups to compare students’ work. (They will have been instructed to select this work during their teaching practice.)

- Ask them to select and assemble work to illustrate two levels – say level 3 and level 5.

- The key to this activity is for them to compare assembled work, not the work of individual students. You want them to argue about standards, the level descriptions, how to arrive at best-fit judgements and the relationship between teacher assessment and day-to-day marking.

GCSE

External (or public) examinations have also witnessed major structural changes. The GCSE replaced both the GCE O-level and CSE in 1986, ending over 20 years of a two-tier examination system at 16. The subsequent introduction of tiered papers for different ‘abilities’ within the GCSE, and the creation of Part 1 GNVQs in 1998, was designed to broaden options at 14, but raised questions about the status and comparability of these assessments. Some time should be spent with a Leisure and Tourism specification, for geography teachers often, rightly, teach this vocational pathway.

Note: It is essential to use awarding body specification documents in training, to familiarise new teachers with the conceptual landscape of external assessment: grade criteria, assessment objectives, weightings, coursework and moderation.

The Pilot GCSE

In 2003 a pilot geography ‘hybrid GCSE’ (see Pilot GCSE project), administered by OCR, was introduced to meet the challenge of broadening post-14 academic and vocational pathways in education generally. Teachers and students have responded enthusiastically to this curriculum development, and it is hoped that it will transmit to other GCSE specifications and refresh the subject. Your student teachers can be active agents in this regard.

The Pilot GCSE is experimenting with linking teacher assessment to external assessment. For this to work, of course, teacher judgements must be both dependable (see Lambert and Lines, 2000) and trustworthy.

Discuss the significance of this with new teachers: do they want to be directly involved in GCSE assessment and curriculum development? What guidance and training would they need?

The GCE A-level

A-level has provided the ‘gold standard’ qualification for university entrance for 50 years. Criticised for over-specialisation, the A-level has so far resisted numerous attempts at fundamental reform. However, most A-levels are already modular rather than linear, and all awarding bodies introduced A- and AS-level geography examinations from September 2000.

The Tomlinson 14-19 Review may mark the beginning of overdue radical change (see the GA’s response). Another useful exercise is to ask student teachers to examine a summary of the Tomlinson proposals and to discuss the place of geography within the Diploma structure.
Planning
Lesson plans should all take account of the assessment. They can be constructed to show an awareness of the programme of study, of the learning objectives and outcomes desired, and an appreciation of how these will be assessed. Opportunities for assessment should be considered from the beginning, not as an afterthought. Unfortunately inspection evidence in geography suggests that assessment is not always sufficiently integrated into planning (Butt and Smith, 1998; Ofsted, 1999) – though this is not necessarily true of new teachers or teachers in initial training!

It is helpful to plan assessment for the short, medium and long term (see Hopkin and Telfer 2000). Teachers can take account of students’ progression against the geography curriculum and plan a range of assessment activities to give students the opportunities to show what they know, understand and can do. Government agencies regularly produce exemplification materials.

High stakes and low stakes
External examinations are often termed ‘high stakes’ summative assessments, in that their results can determine a student’s long-term future. For example, selecting future pathways for students. Examination results can also be easily constructed into league tables and individual school’s and teachers’ performances compared. As a consequence there is a temptation for many teachers to ‘teach to the test’, skewing the educational experience of students. In geography this can mean:

- less discussion and open-ended learning
- fieldwork is only experienced if it directly contributes to examination coursework
- games, role plays and simulations are avoided
- ‘case study’ examples are learned by rote
- photocopied materials may simply be passed to students to learn.

‘Low stakes’ assessments are those made by teachers on a day-to-day basis. They are often formative, designed to support students’ future learning. Their purpose is therefore mainly educational - unlike summative assessments whose purpose is in large part bureaucratic. The point is not to compare each student’s performance with that of his or her peers, but to determine what the next educational steps should be. Low stakes assessments in geography can be extremely varied, ranging from the assessed presentation of a group work task, to a piece of single project work, or to an extended writing exercise.

Norm- and criterion-referenced assessment
The traditional forms of assessment – and those used for external examinations – are generally norm referenced. Students are tested and ascribed a grade related to the marks they have gained. The students’ work is judged in comparison with that of others who have sat the same examination; no pre-determined fixed standard exists for particular grades. The awarding of grades depends on the performance of the whole population of students taking the test.
By contrast in a criterion-referenced assessment the performance of the whole population plays a part. Here the standards of performance are ‘fixed’ into a set of written criterion before the assessment takes place (rather like the driving test). If a particular student matches a particular criteria for performance and attainment he or she will be awarded the grade ascribed to that criterion. Therefore, it does not matter how other students perform - any number of students who meet a given criterion can be awarded the grade which accompanies it.

In practice it is extremely difficult to establish clear criteria in geography, because learning geography is not simply a matter of demonstrating skills and regurgitating knowledge. A good geography test requires students to apply understanding: to imagine that this can be assessed using criteria is like searching for the Holy Grail.

**Conclusion**

Assessment requires effective professional judgement, skill, competence and confidence - especially when one has to adjust earlier judgements about students’ attainments and abilities (Lambert, 1997). Because the interpretation of results gained using various assessment instruments is so complex, it is worthwhile reminding ourselves of their fallibility. This is due partly to ‘user error’, but also to our unrealistic expectations of summative assessment systems.

Our view of assessment is heavily influenced by our view of intelligence. If we believe that students’ abilities are largely fixed at birth, then the role of the teacher is merely to prepare them for assessments whose results are already determined by their genes. If, however, our view is that intelligence is not fixed and immutable then teachers can ‘make a difference’.

As a result, some teachers avoid using the word ‘ability’ when talking about students – or at least to use it carefully. This counteracts the long shadow of the ‘fixed IQ’ that seems to hang over the education service in England. Lumping students together as ‘less able’, for example, is to condemn them unfairly to limited achievement.

**Terminology**

**Certification** - the provision of an award, usually as a paper qualification, for performance in given assessment(s).

**Diagnostic assessment** - assessment which targets particular learning difficulties faced by the student so that future teaching and learning activities can be designed to counter them.

**Evaluation** - the use of assessment information to appraise an aspect of the teaching and learning process. This might be an appraisal of the assessment task itself, of the teaching previously carried out, or of the learning that has resulted.

**Formal assessment** - an internal or external ‘occasion’. Examples would range from a short test or special assessment task to public examinations.

**Formative assessment** - using formally or informally collected information about students’ learning to plan and support future learning.

**Informal assessment** - part of the classroom routine, for example, observing and questioning students as they work, or through discussion.
**Summative assessment** - using assessment information, usually from a variety of tasks, to produce a statement of what a student knows, understands and can do.

**Teacher assessment** – judgement, based on evidence obtained from the continuing process of assessment, through which teachers monitor and record student progress.

The information gathered is used both to support future student learning and eventually to make a summative judgement of student progress. At key stage 3, judgements about national curriculum levels are based on teacher assessment.

*Adapted from Butt, 2000, p. 18.*

**Bibliography**

**Official sources**


ACCAC (2001) Optional assessment materials for geography, levels 1 and 2 at key stage 3. Cardiff: ACCAC.

ACCAC (2001) Optional Assessment Materials for Geography at Key Stage 2. Cardiff: ACCAC.


SCAA (1996b) Consistency in Teacher Assessment: Optional tests and tasks - key stage 3, geography. London: SCAA.


**From the GA**


Butt (pages 17-23) Hopkin (pages 37-45), Hopkin and Telfer (pages 24-35) and Howes (pages 60-9).

**Key texts**

**Journal abstracts**