

A review of literature regarding the validity of coursework and the rationale for its inclusion in the GCSE

Victoria Crisp Research Division

Introduction

The GCSE was introduced in 1988 and is available in a wide range of subjects. GCSEs are assessed mostly by traditional examination, however, in many subjects a percentage of the assessment is via coursework. In the National Criteria Glossary of Terms coursework is defined as 'all types of activity carried out by candidates during their course of study and assessed for examination purposes' (SEC, 1986, p. 1). In practice, coursework takes a wide range of forms: from written reports of fieldwork in geography, to performances and compositions in music and from pieces of art work, to oral contributions during lessons in English. Coursework tends to involve the assessment of a student's work over a period of time. GCSE coursework (in most cases) is assessed by teachers, internally moderated across teachers within schools and then externally moderated by examiners.

Coursework was included in many GCSEs from their introduction to increase the validity of assessment by providing wider evidence of student work and to enhance pupil learning by valuing skills such as critical thinking and independent learning (SEC, 1985). As the Secondary Examinations Council put it 'above all, the assessment of coursework can correspond much more closely to the scale of values in this wider world, where the individual is judged as much by his or her style of working and ability to cooperate with colleagues as by the eventual product' (SEC, 1985, p. 6). Certain types of subject relevant skills cannot be tested via traditional examinations and the inclusion of a coursework unit as part of the relevant GCSE accommodates the assessment of these skills.

There is continuing debate over whether teachers can be trusted to assess their own students. Some argue that teachers' judgements cannot be free from bias whilst others claim that assessment by teachers is the most valid method (as they see a student's work over a period of time) and that teachers' professional judgements should be trusted. Research evidence shows that the validity and reliability of teacher assessment varies and may be related to certain features such as the degree of specification of tasks and criteria (Harlen, 2004), school cultures (Ellis, 1998) and moderation procedures. Experience suggests that in most cases teachers can successfully rank order candidates' work (although some teachers' marking may be more lenient or more severe than others and require adjustment) and the way that coursework assessment is operationalised and standardised makes use of this fact.

The validity and reliability of the assessment of GCSE coursework has come under much discussion since its introduction with the focus of concerns changing over time. At the inception of the GCSE the main threats anticipated were possible unreliability of teacher marking, possible cheating and concern that girls were favoured (see QCA, 2006a). Now, concerns about consistency across similar subjects, fairness and authenticity (including the issues of internet plagiarism and excessive

assistance from others), tasks becoming overly-structured (and hence reducing learning benefits) along with the overall burden on students across subjects, have led to a review of coursework by the Qualifications and Curriculum Authority (QCA). In order to engage with these issues we first need to consider the concepts of validity and reliability.

Validity and reliability

Validity and reliability are central concepts to assessment and describe the confidence we can have in assessment results. Whilst there are slightly different definitions of both reliability and validity, most would agree on the core meanings of these concepts. Reliability is about whether an assessment is repeatable or measures consistently, with a minimum of error. Much attention is given to this issue in assessment development and procedures. The validity of an assessment is about the degree to which it really measures what it purports to measure. Validity and reliability are closely related as a lack of either will result in an assessment that is of little value. In addition, changes to an assessment made to improve validity will often reduce reliability and vice versa.

The traditional view of validity is that there are different kinds of validity: content validity (how appropriate the content of the assessment is as a test of what it aims to assess), construct validity (how well the assessment measures appropriate underlying constructs) and criterion-related validity (how well an assessment relates to actual performance on a specified criterion; this can be predictive or concurrent). In the last few decades most validity theorists have come to consider the construct-content-criterion framework inadequate on the grounds that content and criterion-related validity are actually just examples of evidence that support construct validity. Both Cronbach (1988; 1989) and Messick (1989) consider construct validity the central form. Within this view Messick describes two main threats to construct validity: 'construct under-representation' (the assessment fails to capture important aspects of the construct) and 'construct-irrelevant variance' (capabilities that are irrelevant to the construct are assessed).

Around the same time there was also an emerging view that the concept of validity should be extended to include the consequences of assessment use (Cronbach, 1988; Messick, 1989; Shepard, 1993) specifically with regard to the use of test results, impact on instruction and social consequences. This would include the consideration of whether performance assessment leads to better instructional approaches and does not result in undesirable effects such as narrowing the curriculum (Haertel, 1992). In the climate of both these revisions to the dominant notion of validity, attempts have been made to characterise the types of evidence needed to support construct validity

(e.g. Frederiksen and Collins, 1989; Messick, 1989; Linn, Baker and Dunbar, 1991; Crooks, Kane and Cohen, 1996).

The work of Crooks, Kane and Cohen will be used to provide a structure within which to discuss the validity of GCSE coursework assessment. Crooks, Kane and Cohen's set of criteria has been chosen over those of others as it allows us to focus on the validity that coursework may add as part of a full qualification in comparison to qualifications based only on examinations. In addition, it maps onto key conceptualisations by Messick (1995) and Cronbach (1988) but provides a more practical scaffold for evaluating validity. Crooks, Kane and Cohen (1996) depict assessment validity enquiries as a chain of eight linked stages in order to provide a structure for considering the validity of an assessment. The stages defined are: administration, scoring, aggregation, generalisation, extrapolation, evaluation, decision and impact. For each stage possible threats to validity are exemplified. Crooks, Kane and Cohen suggest that considering possible threats at each stage will allow any 'weak links' to be identified for an assessment.

Validity can be considered a prerequisite to reliability. Crooks, Kane and Cohen (1996) see inter-marker and intra-marker reliability as part of validity because they affect the confidence with which inferences can be made. In the case of coursework, the intention for its use is to improve validity but it may mean greater risks for reliability. Risks to reliability are minimised, at least to some extent, by quality control procedures. However, some teachers initially sympathetic to coursework when the GCSE was introduced were later concerned that the administrative controls put in place to ensure reliability were preventing coursework from being the teacher-led educational experience it should be (Kingdon and Stobart, 1988) and hence limiting the increased validity that coursework was intended to provide.

The validity of GCSE coursework

Although coursework was not a new method of assessment (e.g. it had previously been an optional element of CSEs¹) it was the introduction of GCSE that saw a much increased presence of coursework as part of the assessment culture through its requirement in most subjects. According to Kingdon and Stobart (1988):

... by the time that the GCSE was being introduced, teacher assessment was seen as just another examination technique. Greater understanding of the pros and cons of all techniques had indicated that problems associated with teacher assessment were perhaps no greater than those of other techniques, simply of a different kind. (p. 57)

The reasons for its introduction were mostly about providing a more valid assessment and allowing the assessment of objectives that cannot be assessed by examination, providing complementary assessment of the same objectives, or to assess objectives for which the evidence is ephemeral (SEC, 1986). As the Secondary Examinations Council state, the aim 'should be one of making what is important measurable rather than of making what is measurable important' (SEC, 1985, p. 2).

Despite the apparent advantages of coursework in terms of validity recent concerns such as the new threat from internet plagiarism led the 2005 *14–19 Education and Skills White Paper* (DfES, 2005) to present QCA with a remit to reconsider the value of coursework and address possible concerns. The initial review (QCA, 2005) involved questionnaires

to centres, interviews with teachers, candidates and parents, statistical research and a conference day with examiners. Further work has included a MORI telephone survey of teachers' views (MORI, 2006), a review using this evidence and evidence from QCA monitoring (QCA, 2006a) and an online survey of views on coursework in Maths GCSE (QCA, 2006b).

This section will now use the stages of assessment described by Crooks, Kane and Cohen (1996) to structure discussion of possible improvements to the validity of assessment as a result of including a coursework element within GCSE specifications and possible threats to validity associated with coursework. The stages or links will be considered in reverse order as advised by the authors.

Impact on the student and other participants arising from the assessment processes, interpretations, and decisions

This link in the assessment process as described by Crooks, Kane and Cohen looks at the consequential basis of validity. The direct and indirect impacts of assessment are to be considered along with the effects of experiencing the whole process of assessment. Crooks, Kane and Cohen suggest that threats to validity here can include positive consequences not being achieved or the occurrence of a negative impact of the assessment. The inclusion of coursework in the GCSE was intended to have a positive impact on validity in this respect by providing a number of benefits to learning such as promoting skills of critical thinking, creativity, independent thinking, communication, research and reflection on work (SEC, 1985; SEC, 1986) and allowing helpful feedback from teachers (Ogborn, 1991). Coursework was also intended to be motivating through the realistic sense of audience, the opportunity to negotiate the task and continual assessment (SEC, 1985; SEC, 1986; Ogborn, 1991). In addition, Ogborn (1991) argues that coursework forces teachers to plan courses carefully. In these ways the use of coursework might reduce some threats to validity to do with impact that may exist where assessment consists of examinations alone (e.g. focusing on factual knowledge at the expense of higher order skills). However, if concerns about coursework becoming overly formulaic and predictable in some subjects are well-founded, then coursework may not achieve its intended positive impact. Achieving positive impacts may also be at risk if some students only engage with coursework tasks at a surface level.

Additionally, the heavy workload for teachers and students reported by some constitutes a negative impact of coursework for some of those involved and hence may threaten validity in this respect. In the early days of the GCSE efforts were made to address this concern and various teachers and Local Education Authority professionals investigated and sought to provide advice and good practice ideas based on experience. The main means of controlling the demand of coursework is thought to be to 'ensure that coursework is integrated into the curriculum' (SEC, 1985, p. 8) with tasks arising out of good classroom practice (Cunningham, 1991). Possibilities such as using one piece of coursework to address requirements of more than one subject (Leonard, 1991) or to use methods other than writing were tested but did not become common practice (except for the current overlap in coursework between English and English Literature). Cross-curricular schemes required extra planning from teachers but did reduce student workloads (Leonard, 1991).

It is interesting to note that Scott (1990) found that only a small number of pupils were doing excessive amounts of coursework and other homework. He also reported that the way that pupils reacted to coursework and homework pressure was not related to the amount they actually had to do.

1 CSE was a predecessor of the GCSE.

Decision on actions to be taken in light of judgements

Crooks, Kane and Cohen's 'decision' link is about actions that are taken as a result of judgements, for example, whether a student's score is considered appropriate to admit them to a course. When evaluating the validity of an assessment this stage involves evaluating the merit of the decisions that are taken and whether they are consistent with the information on which they are based and have generally beneficial consequences. One possible threat to validity at this stage according to Crooks, Kane and Cohen would be poor pedagogical decisions. The inclusion of coursework actually gives space for teachers to make good pedagogical decisions. They have more scope to provide useful feedback to students and greater freedom and flexibility within the curriculum, the latter of which was reported by teachers in MORI's survey for QCA (MORI, 2006). However, there is a risk that some teachers may dedicate too much time to coursework at the expense of other areas of study.

Evaluation of the student's performance, forming judgements

This link in the assessment chain is about evaluating what the scores relating to the target domain mean, for example, evaluating what the scores tell us about a student's strengths and weaknesses. Potential threats to validity at this stage can include biased interpretations of test scores (e.g. as a result of a 'halo effect') and poor understanding of the assessment information and its limitations. These issues are the same for GCSE results regardless of whether coursework formed part of the assessment and are hence beyond the scope of the current discussion.

Extrapolation from the assessed domain to a target domain containing all tasks relevant to the proposed interpretation

In the extrapolation link we consider the validity of extrapolating assessment results from the assessed domain to the target domain. This might usually be termed 'generalisability'. According to Crooks, Kane and Cohen, overly constrained assessment conditions would threaten validity in the extrapolation link. This threat to validity is likely to be reduced by the inclusion of a coursework element as part of a qualification.

Another potential threat to validity in terms of extrapolation occurs if parts of the target domain are not assessed or are given minimal weight. This is similar to 'construct under-representation' as described by Messick (1989). The inclusion of coursework in GCSE assessment is likely to improve validity in this respect as it allows types of skills that cannot be assessed by an examination to be evaluated. Improving construct representation was one of the key aims of including coursework in GCSE from the outset.

Avoiding construct under-representation is just as important today as it was when GCSEs were introduced but it seems to be that other threats to validity are currently considered greater concerns and are resulting in changes in the use of coursework.

Generalisation from the particular tasks included in a combined score to the whole domain of similar tasks (the assessed domain)

This link considers the accuracy of generalising from an aggregated score in an assessment to performance in the entire assessed domain (e.g. the entire range of tasks falling within the specification). If the conditions of the assessment vary too much then this can make such generalisations problematic. The term reliability would often be used to describe this issue. With coursework, the conditions do vary somewhat and the tasks

used vary but this may be necessary in order for coursework to achieve its purpose of broadening the skills assessed without becoming so over-defined that the learning benefits are lost and risks of plagiarism are increased. The assessment of only a small sample of student work would also threaten reliability. Coursework can involve just one or two tasks but these are large tasks conducted over a longer period of time so they effectively increase the sample size for a GCSE qualification more than could be achieved using an equivalent exam and hence should help to avoid 'construct under-representation' (Messick, 1989).

Aggregation of the scores on individual tasks to produce one or more combined scores (total score of subscale scores)

Issues under Crooks, Kane and Cohen's aggregation link include aggregating tasks that are too diverse and giving inappropriate weights to different aspects of assessment. Whilst the aggregation of scores from coursework and other examined components to determine GCSE grades could be considered an aggregation of diverse tasks, this is not generally considered a problem for the use of coursework. If anything, it is a strength since a wider range of relevant skills can be assessed.

Scoring of the student's performances on the tasks

With regard to the scoring of an assessment, Crooks, Kane and Cohen suggest consideration of aspects that can reduce the validity of score interpretations and consequent decisions. One potential risk to the validity of an assessment in this link is that scoring might fail to capture important qualities of task performance. As Crooks, Kane and Cohen describe 'attempts to increase rater agreement by using more objective scoring criteria will often lead to a narrowing of the factors included in the scoring, thereby increasing the risk posed by this threat to validity' (p. 272). This is something that needs to be kept in mind in the context of the design of coursework guidance and mark schemes in individual GCSE subjects. Coursework assessment offers an improvement on examinations in that there is less risk of scoring emphasising unimportant but easily rated aspects of student performance. However, whilst it has been argued that providing wider evidence of pupil work through coursework will increase the repeatability of the assessment (SEC, 1985; SEC, 1986), it was always acknowledged that monitoring the marking reliability associated with GCSE coursework assessment would be important. Indeed, many of the negative responses to the introduction of GCSE involved fears that coursework marking would be unreliable and easily open to abuse (Kingdon and Stobart, 1988). Leonard (1991) discusses the 'tension between trusting the professional judgement of teachers and the issue of public confidence in the system of assessment' (p. 10). It is perhaps counter-intuitive to public opinion that teachers can judge their own students without bias.

Some data are available on the reliability of coursework marking. Taylor (1992) asked two moderators to re-mark pieces of coursework in each of GCSE English, maths and history and A-Level psychology and compared the marks given between the two moderators with the mark given by the original moderator. Good correlations between different pairs of moderators were found in each subject (ranging from 0.73 to 0.97). Additionally, Taylor found evidence that there were many more centres that over-marked candidates than under-marked. Wiliam (1996) mentions evidence that in the marking of the 100 percent coursework English GCSE teachers learnt to agree on what grade a piece of coursework was worth but they did not always agree on the aspects of the work that were most significant in making the work worth a particular grade.

It is interesting that Crooks, Kane and Cohen comment when discussing potential marker consistency that 'it is desirable to reduce the extent of such inconsistency, but not at the expense of eliminating or reducing the weight given to important aspects of task performance which can only be assessed through professional judgement' (p. 272).

Administration of the assessment tasks to the student

The conditions under which students take an assessment can impact on the validity of interpretations about the assessment and this link in Crooks, Kane and Cohen's model involves examining the task administration. The use of coursework eases the threat to validity caused by stress in exams and is thought to improve motivation. For example, coursework is thought to be fairer for hard-working pupils who are affected by exam stress and also allows the use of tasks that would cause anxiety in an exam situation (SEC, 1985; SEC, 1986). However, the testing conditions involved in coursework can be dissimilar (Scott, 1990) and clashing deadlines for coursework completion across subjects may cause anxiety for some students.

The threat to validity that seems to be considered most significant currently comes under the category of 'administration' and is about ensuring authenticity of student work. As a result of such concerns coursework is currently being reviewed by the Qualifications and Curriculum Authority (QCA). Concerns relate to plagiarism and excessive assistance from others in particular. The arrival of the internet and increased presence of computers in homes has made the potential for plagiarism greater. Additionally, the level of structure and uniformity of coursework tasks may make plagiarism easier.

Some engagement of parents in their child's coursework is encouraged. QCA's first review report (QCA, 2005) found that nearly two-thirds (63%) of parents helped in some way (e.g. checking spelling and grammar, helping to find an article) and 5% of parents with children taking GCSE admitted to actually drafting some of their child's coursework. The report suggests that there is a lack of awareness that this is not allowed and that there are consequential penalties. Such collusion was always a possibility with coursework but seems to be greater concern now than in the past.

The QCA review (2005) reports that some students admitted trying to download assignments from the internet but not to using them. Some admitted having submitted the work of a sibling or friend as their own. There is also a possibility for inadvertent collusion between peers where part of fieldwork or investigations involves group work or identical tasks.

The QCA (2005) report makes a number of proposals including ensuring that teachers can confirm authenticity, guidelines for teachers and parents on limits of permitted help (these have now been prepared and made available) and giving a higher profile to malpractice. These may help to reduce potential threats to validity in this link.

Strong and weak validity links for coursework

Using Crooks, Kane and Cohen's model we can identify the links where coursework reduces threats to validity compared with examinations alone and links where threats to validity remain for coursework. Coursework has strengths in terms of improving construct representation (extrapolation), the potential for positive effects on learning (impact) and increasing motivation and reducing assessment anxiety (administration). When GCSEs began, the threats to validity that caused concern were possible negative effects in terms of impact due to workload for teachers

(impact) and the potential for biased or inconsistent marking by teachers (scoring). Recently, concerns have shifted towards the issue of authenticating work (administration) and it is this threat to validity, combined with workload issues for students and teachers that seem to be central in driving current changes.

The future of GCSE coursework

As mentioned earlier, the *14–19 Education and Skills White Paper* pointed to concerns about GCSE coursework and gave QCA the remit of addressing certain issues. The 2005 QCA report concluded that the use of coursework needs review in a number of subjects but that it may not be needed in some subjects. A series of reviews were instigated starting with one focussed on mathematics (given that 66% of teachers felt mathematics coursework was problematic) and a MORI study of teachers' views across seven subjects (QCA, 2006a). The QCA has now confirmed that coursework will be dropped from GCSE mathematics from courses beginning in September 2007 and from a number of other subjects (business studies, classical subjects, economics, English literature, geography, history, modern foreign languages, religious studies and social sciences) from courses beginning in September 2008 where they will be replaced with controlled assessments. Controlled assessments are likely to involve tasks being set or approved by the awarding body, conducted under supervised conditions and marked by teachers (QCA, 2007). This would mean a reduction in possible threats to validity in terms of authentication (administration link) and perhaps in terms of marking reliability (scoring link). However, it could have the potential to reduce the validity benefits of coursework in terms of construct representation (extrapolation link) if tasks limited the skills tested, or to reduce validity benefits in terms of impact if tasks became less interesting or overly structured. It is difficult to be sure of the likely effects on validity until the exact nature of controlled assessments is known.

Decisions over changes have been justified by QCA on the basis of three key principles: that the intended learning outcomes in the subject should be critical in determining the appropriate form of assessment, that the most valid (including reliable) form of assessment for a learning outcome should be used so that results are fair and robust and maintain confidence, and that the assessment process should be manageable (QCA, 2006a). It is interesting that the Heads of Department interviewed by MORI (2006) were fairly positive about coursework, particularly in subjects with oral or practical coursework tasks, and nearly all acknowledged the benefits to students. Furthermore, the QCA reviews report a general consensus of the positive impact of coursework on teaching, learning and assessment and that the benefits outweigh the drawbacks (QCA, 2005).

Concerns about internet plagiarism were not as great as might have been expected (82% of teachers disagreed that students used the internet too much) and whilst more than half felt that students in some schools can gain unfair advantage in the current system the most frequently mentioned drawback was the burden of marking coursework. The interviews by MORI found that 66% of teachers were opposed to removing coursework and 51% were strongly opposed to its removal. The MORI interview evidence would not seem to support the decisions that have been made though the controlled assessment proposals might well address teacher concerns that removing coursework would impact on teaching (e.g. lead to less time spent on practical tasks or fieldwork).

It seems that concerns about threats to validity in the administration

link (i.e. authenticity, burden) and concerns about workload seem to be out-weighting possible advantages of coursework to validity in terms of construct representation (extrapolation link) and learning experiences (impact link). However, if the controlled assessments could maintain validity in terms of construct representation and learning experiences as well as reducing threats in relation to administration, then they could provide a more robust overall 'chain' of validity links.

References

- Cronbach, L.J. (1988). Five perspectives on validity argument. In: H. Wainer & H.I. Braun (Eds.), *Test validity*. 3–17. Hillsdale, NJ: Erlbaum.
- Cronbach, L.J. (1989). Construct validity after thirty years. In: R.L. Linn (Ed.), *Intelligence: measurement, the theory and public policy*. 147–171. Urbana: University of Illinois Press.
- Crooks, T.J., Kane, M.T. & Cohen, A.S. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy and Practice*, 3, 3, 265–286.
- Cunningham, B. (1991). Coursework in GCSE English and English literature. In: SEAC (Ed.), *Coursework: learning from GCSE experience: an account of the proceedings of a SEAC conference*. London: Secondary Examinations and Assessment Council.
- DfES (2005). *14–19 Education and Skills White Paper*. London: DfES.
- Ellis, S.W. (1998). *Developing whole-school approaches to curriculum and assessment in secondary schools*. Paper presented at the British Educational Research Association Annual Conference, Queen's University, Belfast.
- Frederiksen, J.R. & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18, 9, 27–32.
- Haertel, E. (1992). Performance measurement. In: *Encyclopedia of educational research*. 6th ed., 984–989. New York: Macmillan.
- Harlen, W. (2004). A systematic review of the evidence of reliability and validity of assessment by teachers used for summative purposes. In: *Research Evidence in Education Library*. London: EPPi – Centre, Social Sciences Research Unit, Institute of Education.
- Kingdon, M. & Stobart, G. (1988). *GCSE examined*. Lewes: Falmer Press.
- Leonard, A. (1991). Case studies of school-based innovation. In: SEAC (Ed.), *Coursework: learning from GCSE experience: an account of the proceedings of a SEAC conference*. London: Secondary Examinations and Assessment Council.
- Linn, R.L., Baker, E.L. & Dunbar, S.B. (1991). Complex, performance-based assessment: expectations and validation criteria. *Educational Researcher*, 20, 8, 15–21.
- Messick, S. (1989). Validity. In: R.L. Linn (Ed.), *Educational measurement*. 3rd ed., 13–103. New York: American Council on Education.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 9, 741–749.
- MORI. (2006). *Teachers' views on GCSE coursework: research study conducted for the QCA*. Available at: www.qca.org.uk/downloads/QCA-06-2735_MORI_report_teacher_views_coursework.pdf (accessed 13.10.06).
- Ogborn, J. (1991). In-course assessment: creating and making best use of opportunities. In: SEAC (Ed.), *Coursework: learning from GCSE experience: an account of the proceedings of a SEAC conference*. London: Secondary Examinations and Assessment Council.
- QCA. (2005). *A review of GCE and GCSE coursework arrangements*. London: Qualifications and Curriculum Authority.
- QCA. (2006a). *A review of GCSE coursework*. London: Qualifications and Curriculum Authority. Available at: www.qca.org.uk/downloads/QCA-06-2736_GCSE_coursework_report-June-2006.pdf (accessed 13.10.06).
- QCA. (2006b). *GCSE mathematics coursework: consultation summary*. London: Qualifications and Curriculum Authority. Available at: www.qca.org.uk/downloads/QCA-06-2737_maths_coursework.pdf (accessed 13.10.06).
- QCA. (2007). *Controlled Assessments*. London: Qualifications and Curriculum Authority. Available at: www.qca.org.uk/qca_115533.aspx (accessed 06.08.07).
- Scott, D. (1990). *Coursework and coursework assessment in the GCSE*. Coventry: University of Warwick, Centre for Educational Development, Appraisal and Research, CEDAR reports, no.6.
- SEC. (1985). *Working paper 2: coursework assessment in GCSE*. London: Secondary Examinations Council.
- SEC. (1986). *Working paper 3: policy and practice in school-based assessment*. London: Secondary Examinations Council.
- Shepard, L.A. (1993). Evaluating test validity. *Review of Research in Education*, 19, 405–450.
- Taylor, M. (1992). *The reliability of judgements made by coursework assessors*. AEB Research Report RAC 577.
- William, D. (1996). Standards in examinations: a matter of trust? *The Curriculum Journal*, 7, 3, 293–306.

ASSESSMENT JUDGEMENTS

School-based assessment in international practice

Martin Johnson Research Division and **Newman Burdett** Cambridge International Examinations

Introduction

This article is based on research and experience in a wide variety of circumstances, educational back drops, social, cultural and political imperatives and, therefore, the proposals and guidelines need to be taken in context; it is impossible to argue whether a Ferrari or a Land Rover is a better car unless you know how it is to be used.

The term 'school-based assessment' (SBA) can conjure up diverse and not necessarily synonymous meanings which often include forms of ongoing and continual classroom assessment of a formative nature.

Sometimes the term is simply used to distinguish localised assessment arrangements from other externally imposed forms of testing. In this article we have defined SBA in a more restricted sense; using it to describe the assessment of coursework. The UK Qualifications and Curriculum Authority (QCA) define coursework as 'any type of assessment of candidate performance made by the school or college in accordance with the specification (or syllabus) of the course of study that contributes to the final grade awarded for a qualification' (QCA, 2005, p.6). QCA go on to identify a number of activities that might be suitable for coursework assessment, and these include: written work and