

Validity issues in the reform of a practical science assessment: An English case study

Neil Wade is a science subject specialist for OCR (Oxford, Cambridge and RSA), a division of Cambridge Assessment, and a member of the Inter-Board Working Group for the Practical Endorsement for GCE A levels in England. (neil.wade@ocr.org.uk)

Ian Abrahams is Professor of Science Education and Head of the School of Education at the University of Lincoln, England. (IAbrahams@lincoln.ac.uk)

Background

Whilst assessments in national examinations are frequently modified, radical change, such as changes to the overall form and style of assessment, is infrequent. However, as part of a wider reform of A-levels (qualifications taken by 18 year olds in England as preparation for university study), the assessment of practical skills and techniques in A-levels in biology, chemistry and physics will change radically from September 2015.

Currently practical skills at A-level are assessed using tasks set externally but marked internally by teachers, referred to as controlled assessments. These tasks contribute to the overall grade awarded for the subject. Central to the reform of the assessment of practical skills is the separation of what we refer to as the direct assessment of practical skills (DAPS), assessed in the classroom by the teacher, from indirect assessment of practical skills (IAPS), assessed in written examinations. Additionally, the change involves giving a separate reported grade, (pass/not-classified), for the practical skills and techniques demonstrated in class, alongside the grade for the written examination. The intention is to increase the validity of the resulting grade, whilst improving the students' competence in specified practical skills and techniques which will be of value either in the work-place or at university.

In this paper we explore the factors which threaten the validity of practical assessments currently used at A-level, and the potential impact of the reforms. In order to do so we trace the development of DAPS and IAPS (Abrahams, Reiss & Sharpe, 2013): their initial conception; their discussion within the Department for Education in England and Ofqual (a non-ministerial government department which regulates qualifications and assessments); and finally the interpretation of those policy requirements by the awarding organisations who provide the A-level examinations (Evans & Wade, 2015).

The Structure of National Assessment in England

The Department for Education (DfE) is responsible for teaching and learning, including the definition of subject content for biology, chemistry and physics at GCE AS and A level, and learning in all maintained (state funded) schools (DfE, 2015). Qualifications, examinations and assessments are separately regulated - including their validity - by a non-ministerial department, Ofqual, the Office of Qualifications and Examinations Regulation (Ofqual, 2015a). The implementation of any changes directed by DfE and Ofqual is the responsibility of the four awarding organisations (AOs) AQA, OCR, Pearson (Edexcel) and Eduqas each of whom offers their own interpretation of the requirements for each subject. Each new specification is submitted to Ofqual for accreditation prior to being made available for schools and colleges (collectively referred to as centres).

Existing qualifications

GCE A levels are taken by 18 year old students following a two year programme of study. Currently these are modular exams with the opportunity for multiple resit. The modules are

predominantly content based with separate controlled assessments for practical skills which have a similar structure for all AOs, a defined task being carried out by all students within a centre on a specific day. The task is completed against a definitive question paper, marked internally by the teacher against a precise mark scheme, before external moderation.

AO dissatisfaction with the existing qualifications

Cambridge Assessment called for alternative approaches in coursework components of exams (Oates, 2013) following research initiated for the international review of the English National Curriculum in 2010 and subsequent analysis of coursework data during 2011. More recently, Watt (2013), citing a paper published by AQA, one of the AOs (Weadon & Stockford, 2010), describes current assessment arrangements as “Inherently unreliable” (p18)

It is accepted that over the duration of a qualification there will be grade inflation as teachers become familiar with the content and style of a qualification. As Stobart (2008) comments:

One of the biggest threats to more effective learning is when schools and teachers, for accountability purposes, gear themselves to get results on tests. This can then narrow down much of teaching and learning to the preparation for the tests of basic skills. (p115)

This can be seen in the skew in the marks for the practical component of a GCE A level qualification in physics (Figure 1): compared to the written papers, the practical component (left) shows a negative skew, with candidates bunched towards the top of the distribution. This indicates that the assessment is not differentiating well between candidates, and provides further evidence to support the questioning of the validity of this form of assessment.

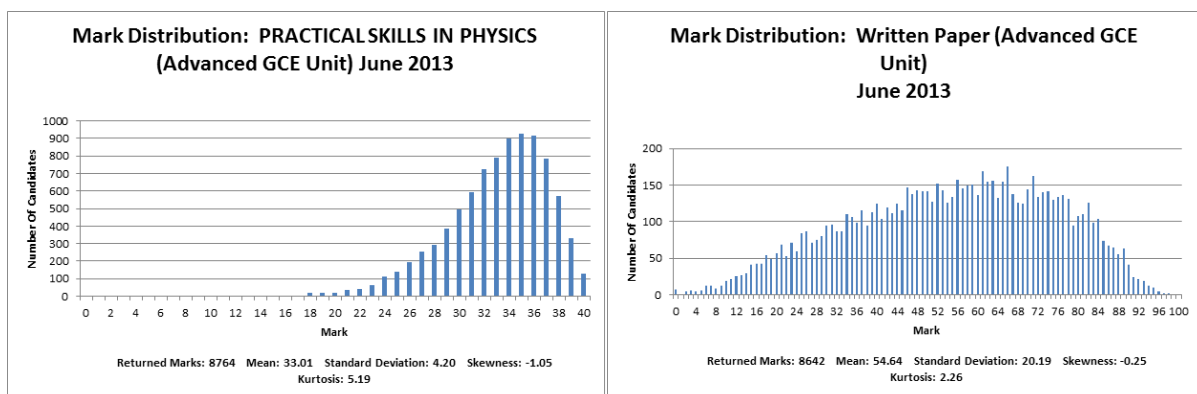


Figure 1. Variation in mark distribution for practical and written assessment

Third parties interested in the reform of practical assessment

Anecdotally universities have observed for some time that students arriving for undergraduate STEM courses had a lack of practical skills. This is supported by a report on the findings of university staff perceptions at a number of English Universities (Grant, 2011). The universities themselves, along with the Gatsby Charitable Trust (that had commissioned its own report – Reiss, Abrahams & Sharpe, 2012), lobbied that there should be a change in the A level system to improve the standard of skills demonstrated by students at the outset of higher education courses. Whilst the universities are considered here as a third party, they are a key stakeholder in the educational process requiring valid and comparable outcomes from the range of examinations available for university entrance, with predictive validity being of paramount importance to enable them to select the best students.

The learned societies representing the sciences (The Institute of Physics, Royal Society of Chemistry and Society of Biology) along with the Royal Society, formed a partnership with the Association for Science Education called SCORE, the science community representing education (Score, 2015a), which also lobbied for changes to the existing qualifications (Score, 2015b).

Observations from teaching and learning

The role of assessment of practical work in science lessons (practical work is a substantial component of what was formerly known in the Science National Curriculum in England as Sc1) has been commented on (Donnelly, 2000) as being primarily used for assessment towards specific examinations rather than for the skills it may provide:

... it appears that Sc1 is most commonly used for purposes of assessment, and more rarely taught, either for the sake of the skills it is intended to promote or as a vehicle for the teaching of scientific content. (There is perhaps an ambiguity here, with teachers indicating that they very often use Sc1 for assessment purposes, rather than that they very often undertake assessment of Sc1.) (p.28)

Indeed, as Nott and Wellington (1999) note:

The skills and processes of investigations are not taught, but experienced, and the conduct of investigations is about summative marks for GCSEs rather than formative assessment to become a competent scientist. In that both pupils and teachers see them as more about getting marks than learning some science, the assessment tail is definitely wagging the science dog. (p. 17)

In a study by Bennett and Kennedy (2001) they reported on “the inadequacies in the current model of assessment of practical skills and abilities, with written examinations questions on practical work examining only a very limited range of abilities” (p.108). Indeed, changes in the way practical work is used in schools has meant, as Toplis and Allen (2012) discuss, that there has been:

... a shift in England and Wales since the 1960's [sic] away from practical work for teaching apparatus handling skills and towards augmentation of knowledge and understanding of substantive concepts, and 21st century UK school science has little to do with the formal assessment of these skills. (p.5)

We believe that as practice in school is largely led by assessment pressure, if there is a desire for teachers to re-focus some of the time spent in doing practical work on developing actual practical skills that will be useful for further study and/or employment, then it is important that students' competencies in such practical skills are formally acknowledged and identified in the assessment process.

The Categorisation of DAPS and IAPS

Whilst Welford, Harlen and Schofield (1985) suggest that “the assessment of practical skills may be possible from pupils' reports or write-ups – **provided that they have actually carried out the practical or investigation prior to putting pen to paper**” (p. 51, bold in original), we would suggest that practical skills are, in many cases, best assessed directly. For example, whilst a conceptual understanding of the topology of knots and manifolds might well be assessed by a written task the most effective means of assessing whether a student is competent in tying their shoe laces is to actually watch them as they attempt to tie them.

In this respect we feel that a useful distinction can be made between what we refer to as the direct assessment of practical skills (DAPS) and indirect assessment of practical skills (IAPS) (Abrahams et al., 2013). The former, DAPS, refers to any form of assessment that requires students, through the physical manipulation of real objects, to directly demonstrate a specific or generic skill in a manner that can be used to determine their level of competence in that skill. An example of this would be if a student was assessed on their skill in actually using an ammeter (in contrast to describing either orally or in writing how they would envisage using an ammeter) and this was determined by requiring them to manipulate a real ammeter, use it within a circuit to take readings, and for these readings to need to be within an acceptable range for the student to be credited.

In contrast, IAPS relates to any form of assessment in which a student's level of competency, again in terms of a specific or generic skill, is indirectly inferred from information they provide, such as reports of the practical work that they undertook or are planning to undertake (e.g. if one is assessing the skill of planning). For example, in indirectly assessing a particular student's competency in the use of an ammeter when the student is working in a group of students who have access to a single ammeter, the marker might be required to make a judgement on the basis of what that student reported they had done (or would do) even if, within the group in which they had undertaken the practical task, the ammeter had (although this might not be reported) only been used by another student.

A common example of the need to use both DAPS and IAPS to best assess both a learner's practical skills (understood broadly to include process skills) and conceptual understanding respectively, and one that we consider provides a useful analogy, is the UK Driving Test. In this example not only does the candidate have directly to demonstrate a sufficient level of competency when actually driving on the road (DAPS) but they must also pass an on-line test to assess their understanding of how to drive a car safely and competently (IAPS).

Table 1 shows a comparison between DAPS and IAPS.

	DAPS	IAPS
What is the principle of the assessment?	A student's competency at the manipulation of real objects is <i>directly</i> determined as they manifest a particular skill	A student's competency at the manipulation of real objects is <i>inferred</i> from their data and/or reports of the practical work they undertook
How is the assessment undertaken?	Observations of students as they undertake a piece of practical work	Marking of student reports written immediately after they undertook a piece of practical work <i>or</i> marking of a written examination paper subsequently taken by students
Advantages	-High validity -Encourages teachers to ensure that students gain expertise at the practical skills that will be assessed	More straightforward for those who are undertaking the assessment
Disadvantages	-More costly -Requires teachers to be trained to undertake the assessment - Has greater moderation requirements	-Lower validity -Less likely to raise students' level of practical skills

There are many cases when the use of IAPS can provide reliable and valid means of assessment. However, their current dominance within summative assessment of practical work in science means that the focus has been directed on to what students know about practical work and how it should, at least in theory, be undertaken rather than on their competency in terms of actually being able to do practical work. This does not, we suggest, seem the best way to assess a student's competency in terms of the practical skills required to make up a buffer solution, use an oscilloscope or prepare a microscope slide. Indeed, over-reliance on IAPS for the assessment of practical work has the potential to lead teachers and students to focus on mastering only 'minds-on' rather than 'hands-on' and 'minds-on' science.

Action taken by government

The government's Education Select Committee work on the administration of exams for 15 – 19 year olds, in particular their consideration of the confidence and credibility with the current system (House of Commons, 2012) identifies a number of key areas of concern such as grade inflation and a loss of confidence amongst Universities and Employers about the rigour of the exams or that the grades accurately reflect the ability of students to apply skills in the work-place. They also note the tension created by the market place in which alternative qualifications are in competition, and refer to this as the "race to the bottom" (paragraph 38).

Ofqual then categorise their concerns in their review of controlled assessment for science (Ofqual, 2013) stating that, despite best efforts, controlled assessment does not always assess what was intended, diverts time from teaching & learning, is arduous to organise and deliver, is delivered inconsistently and does not differentiate well, with the majority of marks clustered between 75% and 95%. Their conclusion is that non-exam assessment should be used only when it is the only way to assess, balances the valid assessment of essential knowledge and skills with sound practice and manageability, fits subject requirements including relative weighting and is not easily distorted by external pressures

The requirements issued

Following public consultations the DfE then issued their requirements (DfE, 2014) for the content of the reformed A levels for first teaching in September 2015. The subject content listed as "knowledge and understanding" (p2 for each science) is intended to constitute sixty percent of the learning outcomes of the specification written by each AO, with the remaining forty percent added to suit to meet the requirements of each AO.

In addition to this, appendix 5 of the document details the requirements for working scientifically. It lists the practical skills assessed by written examination in appendix 5a; the practical skills identified for direct assessment and developed through teaching and learning, which are common to all three sciences, in appendix 5b; the use of apparatus and techniques, as individual lists for each science in appendix 5c.

In 2015 Ofqual then issued specific conditions and requirements relating to the assessment of each subject (Ofqual, 2015c). The requirements for practical assessment are that 15% of the marks in written examinations relate to practical activity, detailed in appendix 5a, whilst students must demonstrate the skills and techniques detailed in appendix 5b and 5c for each science by completing a minimum of twelve practical activities throughout the two year course. Practical skills are separated from the overall grade, and will be reported independently on the basis of pass or not-classified.

Interpretations by AOs

In the process of accreditation by Ofqual (2015b) each AO has interpreted the requirements into a detailed specification, supported by an assessment strategy document. The accreditation process is intended to allow variation in the approach taken, whilst maintaining comparability across specifications and AOs in meeting the requirements.

This has allowed significant variation in the interpretation of the requirements. In its simplest form the definition of twelve required practical activities which incorporate the skills and techniques would satisfy the requirements. However some of those activities may be large group demonstrations, with individual students taking readings, which do not directly address the issues of:

- Increasing the amount of practical activity within science
- Equipping students with the wide range of skills required by higher education
- Accommodating students who are absent for any one practical activity

The model aspires to not only meet those requirements but also allows practical work to be integrated into teaching and learning. This is central to the OCR practical endorsement (Evans & Wade, 2015) but also supported by ‘The Cambridge Approach’ (Cambridge Assessment, 2009) which promotes the move from controlled to performance based assessment. Furthermore it proposes that simulated or naturally occurring activities, which allow assessment against stated standards or assessment objectives, give that assessment greater validity. The move from assessment with very specific tasks on defined dates (the current model) to the ability to assess continuously, through all the opportunities of practical work integrated into teaching and learning has the potential to improve the validity of assessing practical skills.

The next stage of development involved looking at the wide range of reasons that practical work is used in teaching science such as illustrating and demonstrating physical phenomena, familiarisation with basic apparatus and techniques, data gathering and analysis, introducing scientific methods and designing and reporting a scientific investigation (Main, 2014).

Each of these is a valid use of practical activity and the intention is that they all provide opportunities for students to demonstrate practical skills and techniques. The assessment of such skills is then embedded in teaching and learning. One benefit of this approach is the potential to save time by avoiding the perceived need to practice each assessment, as in the current system, before then delivering the formal assessment, which we calculate could be as much as eighteen hours of contact time in a year.

We can also consider that other factors will affect the style of teaching used in any centre, aside from the personal preferences of the teacher. Class size (5 or 25), lesson duration (50 minutes or two hours), laboratory facilities (shared labs or lessons in classrooms), equipment available and the level of technician support may all impact on the scheme of work developed by the centre, a notion supported by point 69 of the Ofsted report “Success in Science” quoted in Wellington and Ireson (2012, p10). The perceived benefit being that each centre can choose to use practical activities matching their circumstances.

The result is that OCR rather than identifying twelve specific practical activities have identified twelve practical activity groups within each specification. Within each practical activity group there are three alternative practical activities. Centres may then select a minimum of twelve activities which cover the skills and techniques (Evans & Wade, 2015).

In practice the majority of centres already carry out more than this number within their current schemes of work.

The cross board working group

Whilst the distinctive interpretation of the endorsement was allowed within the process of reform, it was a requirement that there should be a common form of assessment across all four AOs. To facilitate this, an inter-board working group was established. One of the first issues addressed was to define the common practical assessment criteria (CPAC). These statements are now embodied in the ‘GCE Subject Level Conditions and Requirements for Science (Biology, Chemistry, Physics) and Certificate Requirements’ (Ofqual, 2015c) with overarching statements:

1. Follows written procedures
2. Applies investigative approaches and methods when using instruments and equipment
3. Safely uses a range of practical equipment and materials
4. Makes and records observations
5. Researches, references and reports

The specification statements from appendices 5a, b and c along with the CPAC and wording detailing the requirement that all students demonstrate the required skills and techniques consistently and routinely are incorporated word for word in each AO’s specifications.

The trial

Having determined the common requirements it was important to ascertain that these would then be practicable and manageable. The inter-board group established a trial with the remit as defined in the summary of the report (Inter-board working group, 2014a) which were: to investigate how far the CPAC will enable teachers to make accurate and consistent judgements about learners’ practical competencies and to suggest any required amendments to support this; to consider the support and guidance that would be necessary for schools and colleges nationally to apply the CPAC effectively; to consider arrangements for the conduct and monitoring of practicals for schools, colleges and awarding bodies and make an assessment as to whether these arrangements would be manageable and scalable; to consider appropriate and proportionate arrangements for teachers and learners to record practical activities and attainments; to engage the wider Science community in the trialling process and to make recommendations about appropriate malpractice sanctions.

The trials incorporated centres to cover extremes of cohort size, type of establishment and geographical location. Participants met at the start of the trial to be briefed on the practical endorsement. They were given an open ended assignment to carry out practical work which occurred naturally in their schemes of work, to assess students against the CPAC and to consider those aspects detailed above. During that time monitors from each AO visited the centres and undertook what they considered appropriate observations to ensure the participation in the practical endorsement and interpretation of the CPAC. Teachers and monitors fed back to the inter-board group on their experiences and this informed the on-going development of the practical endorsement.

The conclusions from the trials (Inter-board working group, 2014a) were that: teachers will be able to use the CPAC effectively; AOs will be able to monitor the system effectively and arrangements for the conduct of practicals and for monitoring teachers’ assessment will be scalable. It was noted that further guidance and support for teachers will be required,

particularly, in relation to the investigative nature of CPAC 2 and the new requirements for research and referencing in CPAC 5, which are competencies not explicitly assessed under the current arrangements of practical examinations. The inter-board working group concluded that the new arrangements offer significant opportunities to enhance learners' experiences of the development of practical skills. In addition to the written report, a meeting in London to present its findings to an assembled group incorporating the learned societies, foundations such as Gatsby and Wellcome, Ofqual and Ofsted fulfilling the requirement to engage the wider science community. This feedback is summarised in a video (Inter-board working group for A level science practicals, 2014b)

Implementation

The feedback from the trial allowed the minimum documentary requirements to be established (Inter-board working group for A level science practicals, 2015):

- documented plans to carry out sufficient practical activities which meet the requirements of the CPAC, incorporating skills and techniques detailed, over the course of the A level
- a record of each practical activity undertaken and the date when this was completed
- a record of student attendance
- a record of the criteria being assessed in that practical activity
- a record of which student met the criteria and which did not
- student work showing evidence required for the particular task with date
- any associated materials provided for the practical activity e.g. written instructions given.

Monitoring

The trial also established that the monitoring process was in no way comparable to the existing moderation of student practical assessments. The current tasks are carried out under controlled conditions and marked against a mark scheme by the teacher. Only around two or three marks out of a total of 40 apply to DAPS, the remainder being allocated to IAPS such as drawing graphs, interpreting results and evaluation. Moderation is used to achieve comparability across centres by the adjustment of individual marks. In contrast, monitoring sets out to establish that the centre is providing the opportunity for each student to demonstrate the range of skills and techniques listed and to carry out the minimum of twelve practical activities. The judgements of the practical endorsement are made by teachers in the classroom and *cannot* be moderated. To improve the validity of assessment students' records can be correlated against the teacher records, substantiating that work and skills recorded by the teacher are noted in the student record. Teacher judgement can be assessed in conversation during the observation of practical work and discussion with students completing this process. The result is designed to allow a holistic judgement to be made as to whether or not the centre is complying with the requirements of the practical endorsement. At no point is an individual student's work being judged, or is their grade dependent upon the outcome of the monitoring visit.

Reflections and scope for further study

Whilst the changes to the assessment of practical skills described above are designed to increase the level of student competency in a range of practical skills the effectiveness of this approach has yet to be evaluated. In evaluating the impact we will look at the provision of practical activities in science teaching in schools, to counter the argument that the removal of controlled assessment will lead teachers to drop practical work (Stacey, 2015). If, as we believe, this reform gives the potential for more and more varied practical work, then we

would also expect to see an impact on learning and on the practical skills demonstrated by students. Finally there is the evaluation of the validity of the outcomes for employers and higher education who we would expect to identify the impact in the longer term.

OCR, amongst a number of organisations, has started a medium term survey of the amount and type of practical work currently undertaken in centres. Initial data is collected and will be compared against a second survey after the reformed assessments are in place.

The impact on learning and student outcomes is inextricably linked to the outcomes for employers and higher education. We envisage using a comparative study in which students currently completing the existing A level programme, as well as those in the future having completed the revised programme, will be asked to complete an external measure of practical skill competency; such a measure could be based on employer or university assessments. Likewise it would be possible for students who had completed an apprenticeship training programme to undertake some of the practical tasks associated with the new A level as a means of evaluating the extent to which students assessed as competent in specific practical skills on one programme are able to demonstrate a comparable level of competence in that same skill using a different form of assessment of that skill.

References

- Abrahams, I., Reiss, M., & Sharpe, R. (2013). The assessment of practical work in school science. *Studies in Science Education*, 49(2), 209-251.
- Bennett, J., & Kennedy, D. (2001). Practical work at the upper high school level: The evaluation of a new model of assessment. *International Journal of Science Education*, 23(1), 97-110.
- Cambridge Assessment (2009). *The Cambridge Approach, principles for designing, administering and evaluating assessment*. Retrieved from: <http://www.cambridgeassessment.org.uk/Images/109848-cambridge-approach.pdf>
- Department for Education (DfE) (2014). *GCE AS and A level subject content for science*, Ref: DFE-00356-2014. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/446829/A_level_science_subject_content.pdf
- Department for Education (DfE) (2015). *Department for Education, About us*. Retrieved from: <https://www.gov.uk/government/organisations/department-for-education/about>
- Donnelly, J. (2000). Secondary science teaching under the National Curriculum. *School Science Review*, 81(296), 27-35.
- Evans, S. & Wade, N. (2015). Endorsing the practical endorsement? OCR's approach to practical assessment in science A-levels, *School Science Review*, 96(357), 59-68
- Grant, L. (2011). *Lab skills of new undergraduates*. London: Gatsby Charitable Foundation. Retrieved from: <http://www.gatsby.org.uk/uploads/education/reports/pdf/russell-group-survey-of-lab-skills-report-laura-grant-may-2011.pdf>
- House of Commons (2012). *The administration of examinations for 15-19 year olds in England; Confidence and Credibility: Key issues with the current system*". Retrieved from: <http://www.publications.parliament.uk/pa/cm201213/cmselect/cmeduc/141/14105.htm>
- Inter-board working group for A level science practicals (2014a). *Summary of cross-board trialling of the A level science practical endorsement*. Retrieved from: <http://www.ocr.org.uk/Images/202103-summary-of-cross-board-trialling-of-the-a-level-science-practical-endorsement.pdf>
- Inter-board working group for A level science practicals (2014b) *A level science practicals, making them truly experimental*. Retrieved from:

https://www.youtube.com/watch?list=PLtzR6sheDAMHxz4paXo80Q7ezJJb_C7WH&feature=player_embedded&v=HI0TFzjVafM

Inter-board working group for A level science practicals (2015). *Monitoring of the practical endorsement for A level biology, chemistry and physics for first teaching September 2015*. Retrieved from: <http://www.ocr.org.uk/Images/230006-letter-to-centres-monitoring-of-the-practical-endorsement-from-september-2015.pdf.pdf>

Main, P. (2014). Thinking like a physicist, design criteria for a physics curriculum, *School Science Review*, 95(352), 46–52

Nott, M., & Wellington, J. (1999). The state we're in: Issues in key stage 3 and 4 science. *School Science Review*, 81(294), 13-18.

Oates, T. (2013). *Radical solutions in demanding times: Alternative approaches for appropriate placing of 'coursework components' in GCSE examinations*. Retrieved from: <http://www.cambridgeassessment.org.uk/Images/145830-coursework-radical-solutions-in-demanding-times-.pdf>

Ofqual (2013). *Review of controlled assessment in GCSEs*, Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/377903/2013-06-11-review-of-controlled-assessment-in-GCSEs.pdf

Ofqual, (2015a). *About us*. Retrieved from: <https://www.gov.uk/government/organisations/ofqual/about>

Ofqual (2015b). *Apply to have your qualifications regulated*. Retrieved from: <https://www.gov.uk/apply-to-have-your-qualifications-regulated#who-and-what-ofqual-regulates>

Ofqual (2015c). *GCE Subject level conditions and requirements for science (biology, chemistry, physics) and certificate requirements*. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/447147/2015-07-20-gce-subject-level-conditions-and-requirements-for-science-and-certificate-requirements.pdf

Reiss, M., Abrahams, I., & Sharpe, R. (2012). *Improving the assessment of practical work in school science*. Retrieved from: <http://www.gatsby.org.uk/~media/Files/Education/Gatsby%20%20Improving%20the%20assessment%20of%20practical%20work%20in%20school%20science.ashx>.

SCORE (2015a). *SCORE the science community representing education*. Retrieved from: <http://www.score-education.org/home>

SCORE (2015b). *SCORE consultation responses*. Retrieved from: <http://www.score-education.org/news/consultation-responses>

Stacey, G. (2015). *A radical experiment to end science practicals? That's just not true*, Retrieved from: <http://www.theguardian.com/teacher-network/2015/feb/04/experiment-science-practicals-ofqual>

Stobart, G. (2008). *Testing times: The uses and abuses of assessment*, Abingdon: Routledge
Watts, A. (2013). *The assessment of practical science: A literature review*. Cambridge: Cambridge Assessment,

Weadon, C. & Stockford, I. (2010). *Classification accuracy and consistency in GCSE and A level examinations offered by the Assessment and Qualifications Alliance (AQA) November 2008 to June 2009*. Retrieved from: <https://cerp.aqa.org.uk/research-library/classification-accuracy-and-consistency-gcse-and-level-examinations-offered-assessment-and-qualifications-alliance-aqa-november-2008-june-2009>

Welford, G., Harlen, W. & Schofield, B. (1985). *Practical testing at ages 11, 13 and 15* London: Department of Education and Science.

Wellington, J. and Ireson, G. (2012). *Science Learning, Science Teaching* (3rd ed.), Abingdon: Routledge