

KEY STAGE 2 TESTING AND ACCOUNTABILITY REVIEW

**Submission from Professor Robin Alexander,
Ms Alison Peacock and Professor Wynne Harlen
on behalf of the Cambridge Primary Review**

INTRODUCTION

The Cambridge Primary Review (CPR) has been invited to give both written and oral evidence to the KS2 testing and accountability review chaired by Lord Bew. The CPR devoted a good deal of attention to testing, assessment and wider issues of curriculum, standards and accountability, and its findings are already in the public domain. Since October 2009, when the CPR's final report was published, little has changed in respect of either the national assessment regime or evidence about it, so it makes sense to restate the CPR's position as published rather than strive to construct an alternative text which would merely paraphrase what has already been refined for publication and has been agreed by the final report's authors and the CPR's advisory committee.¹ We are therefore providing:

- a re-statement of the CPR's published position on assessment and contingent matters, extrapolated from the formal conclusions and recommendations of its final report;
- some additional points in light of developments since then;
- as appendices to this statement, copies of the relevant sections of the final report and three of its commissioned research surveys.²

This may seem an oblique way to address the eight issues listed in the assessment review's published remit. However, we believe that the CPR has attended to most or all of these issues, though we note later in our submission that some of them make rather questionable assumptions which we hope the review will test rather than accept.

Since one of the CPR's key recommendations was that the government should launch precisely the kind of review of assessment that Lord Bew is now leading, and that such a review should assemble the best expertise available to move forward from our analysis of current problems to solutions, the CPR evidence identifies principles on which an improved

¹ The lead editor of the CPR final report was Robin Alexander and its authors were Robin Alexander, Michael Armstrong, Julia Flutter, Linda Hargreaves, David Harrison, Wynne Harlen, Elizabeth Hartley-Brewer, Ruth Kershner, John MacBeath, Berry Mayall, Stephanie Northen, Gillian Pugh, Colin Richards and David Utting. The CPR Advisory Committee comprised Gillian Pugh (chair), Robin Alexander, Dawn Austwick, Patricia Clark, Christina Coker, Kevan Collins, Sheila Dainton, Bernadette Duffy, Kate Froom, David Hargreaves, Elizabeth Hartley-Brewer, Diane Hofkins, Hilary Hodgson, Anna House, James Hughes-Hallett, Pat Jefferson, Richard Margrave, Melody Moran, Stephen Pisano, Andrew Pollard and Sue Tite, with Usha Sahni as Ofsted Assessor.

² Alexander, R.J. (ed) (2010) *Children, their World, their Education: final report and recommendations of the Cambridge Primary Review*, Abingdon: Routledge, pp 311-342 and 471-474.
Tymms, P and Merell, C. (2010) 'Standards and quality in English primary schools over time: the national evidence' in Alexander R.J. with Doddington, C., Gray, J., Hargreaves, L. and Kershner, R. (eds) *The Cambridge Primary Review Research Surveys*, Abingdon: Routledge, pp 435-460;
Whetton, C., Ruddock, G. and Twist L. (2010) 'Standards in English primary education: the international national evidence' in Alexander R.J. with Doddington, C., Gray, J., Hargreaves, L. and Kershner, R. (eds) *The Cambridge Primary Review Research Surveys*, Abingdon: Routledge, pp 461-483;
Harlen, W. (2010) 'The quality of learning: assessment alternatives for primary education' in Alexander R.J. with Doddington, C., Gray, J., Hargreaves, L. and Kershner, R. (eds) *The Cambridge Primary Review Research Surveys*, Abingdon: Routledge, pp 484-520.

system might be based rather than detailed structures and procedures. However, we have proposals on the latter which we hope to discuss with the review panel at the oral session on 27 January 2011.

We should add, by way of further validation for what follows, that the CPR is the biggest and most comprehensive enquiry into English primary education since the Plowden report of 1967. Politically and financially independent, its launch in 2006 was preceded by two years of planning and consultation. Between 2007 and 2009 it published 31 interim reports. These were followed, in October 2009, by its final report. The CPR then embarked, with further support from its sponsors (Esmée Fairbairn Foundation) on a three-year programme of dissemination, policy engagement and network-building which will run, in the first instance, to 2012. Its evidence base includes: over 1000 formal written submissions from a wide range of organisations and individuals both inside and outside education, most of them substantial documents (average length 30 pages); 250 local and national ‘soundings’ and other meetings with children, teachers, parents, community representatives, local authorities, national organisations, teaching unions, non-statutory public bodies, opposition parties and government; thousands of emails; 28 comprehensive reviews of published national and international research and other published evidence relating to the review’s ten themes, commissioned from 66 leading academics in 20 universities; and a re-assessment of official demographic, financial, performance and other data. Altogether, in addition to the evidence from the submissions, soundings and official data searches, over 4000 published sources were consulted.

Assessment and testing were prominent aspects of the CPR’s own remit, but its scope meant that these could be linked, as they should be, to questions about aims, curriculum, pedagogy and accountability. One of the problems of much of the assessment literature is that it tends to concentrate on the technicalities of assessment in isolation from these vital contingent issues. We hope that the review team will strive to embed what it concludes and proposes in the matters of educational purpose, principle and substance which give meaning and point to assessment procedures.

THE PUBLISHED POSITION OF THE CAMBRIDGE PRIMARY REVIEW ON ASSESSMENT, TESTING, STANDARDS AND ACCOUNTABILITY IN PRIMARY SCHOOLS

We quote now, with omissions marked by ellipses, the relevant conclusions and recommendations from chapter 24 of the CPR’s final report.³

THE CURRICULUM

38. As children move through the primary phase, their statutory entitlement to a broad and balanced education is increasingly but needlessly compromised by a ‘standards’ agenda which combines high stakes testing and the national strategies’ exclusive focus on literacy and numeracy. It is regrettable that the Rose Review’s remit excluded examination of these issues. The most conspicuous casualties are the arts, the humanities and those kinds of learning in all subjects which require time for talking, problem-solving and the extended exploration of ideas. Memorisation and recall have come to be valued more than understanding and enquiry, and transmission of information more than the pursuit of knowledge in its fuller sense. Worryingly, primary science, which was one of the success stories of the national curriculum’s first decade, has also been squeezed by the national strategies, retaining its albeit reduced place only because it was tested at the end of key stage 2 (from 2010 this ceases too). Science is far too important to both a balanced education and the nation’s future to be allowed to decline in this way.
39. Fuelling this loss of entitlement has been a policy-led belief that curriculum breadth is incompatible with the pursuit of standards in ‘the basics’, and that if anything gives way it must

³ CPR final report, p 493 and pp 496-500.

be breadth. However, evidence going back many decades, including reports from HMI and Ofsted, consistently shows this belief to be unfounded. Standards and breadth are often positively related, and high-performing schools achieve both. This is one of several modern manifestations of the historic divide between 'the basics' (protected) and the rest of the curriculum (viewed as dispensable). Now recognised by many contributors to the Review as a threat to standards as well as entitlement, this split is exacerbated by the relative neglect of the non-core curriculum in initial teacher training, school inspection and professional development. This produces a primary curriculum which, as Ofsted has acknowledged, is often two-tier in terms of quality as well as time.

...

51. National assessment should be reformed (see next section) so that it does its job without compromising children's legal entitlement to a broad and balanced curriculum.

...

ASSESSMENT AND TESTING

62. The world of assessment has changed radically since the time of Plowden. In 1967, the discourse was dominated by assessment for selection, and was mostly concerned with the pros and cons of IQ tests. In 2009, while testing is very much to the fore, there is recognition that assessment has purposes other than selection, particularly for helping learning, and that it can be conducted in many different ways. The government's publication of an assessment for learning (AfL) strategy in 2008, despite the proliferation of tests and targets, recognises the role of assessment in learning and of teachers' judgements in assessing pupils' progress. However, questions have been raised about the interpretation of AfL in the context of current policy, and it is important that the research-informed principles enunciated by the Assessment Reform Group remain to the fore.⁴

63. That said, primary pupils in England are tested more frequently and at an earlier age than in most other countries, and in public and political discussion testing is frequently equated with assessment. This is a serious error – linguistically, technically and educationally. It generates excessive faith in the validity, power and outcomes of tests and diminishes the use of other kinds of assessment which have greater diagnostic and pedagogical value. Testing is just one method of assessment among several.

64. It is often claimed in defence of national tests that they raise standards. In fact, at best the impact of national tests on standards is oblique. The prospect of testing, especially high-stakes testing undertaken in the public arena, forces teachers, pupils and parents to concentrate their attention on those areas of learning to be tested, too often to the exclusion of other activities of considerable educational importance. It is this intensity of focus, and anxiety about the results and their consequences, which make the initial difference to test scores. But it is essentially a halo effect, and it does not last; for it is not testing which raises standards but *good teaching*. The point is obvious but needs to be underlined. Conversely, if testing distorts teaching and the curriculum, as evidence from the Review and elsewhere shows that it does, it may actually depress standards properly defined.

65. Contrary to some claims about testing, it produces results which have lower validity and reliability than is generally assumed. Another myth about testing is that it objectively compensates for teachers' over-favourable judgements about their pupils. In fact, the evidence shows that teachers' ratings of their pupils' attainment, based on a far wider range of evidence, are likely to be lower than their test scores.

66. Use of SAT results to evaluate teachers, schools and local authorities puts pressure on teachers which is transferred to pupils to the detriment of their learning experiences. The

⁴ Assessment Reform Group (2002) *Assessment for Learning: 10 Principles*, Cambridge: University of Cambridge Faculty of Education; Mansell, W. James, M. and the Assessment Reform Group (2009) *Assessment in Schools. Fit for purpose? A commentary by the Teaching and Learning Research Programme*, London: ESRC/TLRP.

process also places heavy and perhaps excessive demands on teachers' and pupils' time, and on local and national resources. What others have called the fiasco of the marking of the 2008 SATs – which left thousands of papers unmarked even after the results had been announced – raises serious questions about the cost-effectiveness and feasibility of the current system, as well as its reliability and validity.

67. Test results are not the best source of data for the multiple functions they perform – measuring pupils' attainment, school and teacher accountability and national monitoring. Despite government claims to the contrary, the use of aggregated test results as a basis for evaluating schools does not provide a fair picture, even when the disputed 'contextual value-added' scores are used. This high-stakes use of test results leads to practices that not only have negative impact on pupils but fail to provide valid information, being based on what can be assessed in time-limited written tests in at most three subjects. The use of the same data for national monitoring also means that we have extremely limited information, collected under stressful conditions, which provides little useful data about national levels of performance and even less about how to improve them. The aggregation of SAT results for monitoring national levels of performance fails to reflect achievements in the full range of the curriculum. ...
- 68. There is an urgent need for a thorough reform of the assessment system in England, going well beyond the May 2009 report of the DCSF 'expert group', to provide a coherent set of practices and procedures suiting the aims of education in the 21st century and to meet the needs for information about the performance of individual pupils, schools, local authorities and the system as a whole. At the heart of this should be the use of assessment to help learning, leading to the development of lifelong learners. This should be supported by a system for summarising, reporting and accrediting children's performance that provides information about all aspects of learning. Separate systems are also required for the external evaluation of schools and for monitoring national standards of performance.**
- 69. No single assessment procedure, including statutory assessment, should be expected to perform both formative and summative functions. Assessment for learning should be uncoupled from assessment for accountability.**
- 70. Children's learning across all aspects of the curriculum, including their developing capacity to learn, should be assessed formatively throughout the primary phase and summatively before transfer to secondary school. This is not straightforward technically, and on the basis of past experience the dangers of a simplistic and reductionist approach are all too evident. Moving to valid, reliable and properly moderated procedures for a broader approach to assessment will require careful research and deliberation.**
- 71. The Review fully accepts the need for summative assessment at the point where pupils move from primary to secondary education, and tests have a place in this process. But while the assessment of literacy and numeracy is essential, a broader, more innovative approach to summative assessment is needed if children's achievements and attainments across the curriculum are to be properly recognised and parents, teachers and children themselves are to have the vital information they need to guide subsequent decisions and choices. Work is now urgently needed on the development of a comprehensive and coherent framework of summative assessment that can be administered unobtrusively and with minimum disruption towards the end of the primary phase.**
- 72. The use of the results of statutory assessment at key stage 2 for monitoring national performance in primary education should be replaced by sample testing using a bank of varied items covering the curriculum as a whole.**
- 73. The practice of publishing primary school performance tables (now known as primary school achievement and attainment tables) based on the results of statutory**

assessment in English and mathematics at the end of key stage 2 should be abandoned.

...

QUALITY, STANDARDS AND ACCOUNTABILITY

75. The evidence on whether standards in primary education have improved is unsafe. At its heart are two areas of difficulty: the validity and reliability of the chosen measures and procedures; and the historical tendency to treat test scores in limited aspects of literacy and numeracy as proxies for standards in education as a whole.
76. At the national level, the assumption that aggregating individual pupils' test results in only three subjects enables trends in attained standards to be identified is problematic. Although the statistics can be computed, their meaning in terms of changes in attainment are brought into question by the limited range of what is tested, by limitations in test technology and by the impact of using the results for high-stakes judgements. We are left with little sound information about whether pupils' attained standards have changed.
77. Subject to these substantial caveats, analysis of national test scores and international achievement surveys appears to show that standards of tested attainment in primary education have been fairly stable over the short period that usable data have been available, with some changes up or down. Pupils' attitudes to their learning in the tested areas are generally positive (though, as is generally found internationally, it appears to decline as pupils approach the end of primary education). There have been modest improvements in primary mathematics standards, especially since 1995, though different datasets tell different stories. The international data from 2001 show high standards in reading among English pupils by comparison with those from other countries, though the more recent data (from 2006 onwards) suggest that the 2001 results may have been misleading. England appears to be above the international average but not exceptionally so. The international data also show considerable improvements in primary science by comparison with other countries, though there have been methodological reservations about the studies in question.
78. However, gains in reading skills may sometimes have been at the expense of pupils' enjoyment of reading. Similarly, there is some evidence of an increase in test-induced stress among primary pupils, especially at key stage 2, and much firmer evidence of pressure on their teachers. The primary curriculum has narrowed in direct response to the perceived demands of the testing regime and the national strategies, to the extent that in many schools children's statutory entitlement to a broad and balanced curriculum has been seriously compromised; yet the national strategies have had, of themselves, a less pronounced impact on reading standards than might have been expected from the level of investment. The historically wide gap between high and low attaining pupils in reading, mathematics and science has persisted: it is already evident at a very young age and widens as children move through the primary phase. There is no reliable evidence on national standards in areas of children's learning outside those aspects of literacy, numeracy and science which have been tested, other than that in many schools such learning appears to have been compromised by the standards drive itself.
79. Schools acknowledge the importance of being held accountable for their work and accept the need for periodic inspection. Ofsted produces useful annual reports on the condition of the system as a whole and surveys on particular issues, on many of which the Review has drawn to its considerable benefit. The collation of evidence from inspections can be used to provide a reasonably valid, if partial, assessment of the quality of English primary education nationally *at a particular time*, assuming that the Ofsted criteria and procedures are accepted. However, Ofsted's school inspection procedures attract a good deal of criticism in relation to their validity, reliability and impact; and because of frequent changes to inspection criteria and procedures, allied to the subjective nature of the process, it is much more difficult to say with confidence whether the overall quality of primary education has improved, deteriorated or remained the same over time. The same difficulty attends Ofsted inspections of individual

schools. Such judgements are compromised by the successive changes Ofsted has instituted in inspection criteria and methodology and by its employment of different teams from one inspection to the next of the same school. Temporal comparisons and claims about long-term trends based on Ofsted data are thus highly problematic.

80. **Teachers and schools can and should have a greater role in the assessment of their pupils and in the evaluation of their provision for learning. In the case of pupil assessment, there is an overwhelming case for extending the range of aspects of attainment that are included in reporting attained standards and in identifying the standards to aim for. At present the pupil attainment data reflect only a small part of the curriculum and within that only aspects which are easily measured by written tests. Greater use of information that teachers can collect as part of their teaching can help learning and, suitably moderated, can provide information which is a better reflection of performance across the full range of the curriculum. Similarly there is a strong case for moderated school self-evaluation across the full range of provision. Such evaluation should help the school's own improvement agenda and not simply be instituted to meet Ofsted requirements.**
81. **Current notions of 'standards' and 'quality' should be replaced by a more comprehensive framework which relates to the entirety of what a school does and how it performs. The Review's proposed statement of aims for primary education might provide the overall criteria for progress and success, combined with appropriate indicators for each of the proposed new aims and curriculum domains. However, we warn against moving from indicators of what can fairly be observed and judged to so-called measures of what cannot in fact be measured.**
82. **Monitoring the performance of the national system of primary education is technically challenging and requires a form of data collection which is different from that which is optimal to promote school and classroom improvement. New provision is recommended, using robust national sampling and building on the lessons of the Assessment of Performance Unit and work in Scotland.**
83. **A new model for school inspection should be explored, with a substantially increased focus on classroom practice, pupil learning and the curriculum as a whole, and within a framework of accountability which directly reinforces processes of school improvement.**
84. **Every effort should be made, at school, local and national levels, to ensure that curriculum, pedagogy, assessment, inspection and teacher training all pull in the same direction and are clearly informed by educational aims and procedural principles such as those proposed in chapter 12.**
85. **We take it as axiomatic that in a public system of education teachers and schools should be fully accountable to parents, children, government and the electorate for what they do. We reject any suggestion that our proposals for the reform of assessment and inspection imply otherwise. For us, the issue is not *whether* schools should be accountable, but *for what and by what means*, and the evidence shows that current approaches are in certain respects unsatisfactory. By insisting on a concept of standards which extends across the full curriculum rather than part of it, we are strengthening rather than weakening school accountability. It is no less important that others involved in primary education, including central and local government, are fully accountable for their part in the process. When responsibilities are shared, accountability should be shared too in order that the precise cause of problems can be speedily and accurately diagnosed and appropriate remedial action can be taken.**

ADDITIONAL POINTS

The CPR's conclusions and recommendations on assessment received widespread support during the period of dissemination which followed the final report's publication. On that basis, assessment reform was listed as one of the eleven policy priorities which were commended to all political parties just before the general election in May 2010.⁵ Though the previous government dismissed the CPR's final report unread, it later changed its stance and set in train a programme of regular in-depth discussions about the report's implications between the CPR, DfE officials and ministers. This was implemented by the new government and is still in progress.

Apart from agreeing to set up an assessment review broadly in line with the CPR's recommendations, the other important development has been on the curriculum front. As noted above, the previous government's Rose review of the primary curriculum explicitly kept curriculum and assessment apart, even though they are necessarily related and the CPR and other studies had shown that existing assessment procedures were compromising primary pupils' statutory entitlement to curriculum breadth, balance and quality. We are encouraged that the coalition government has accepted the interrelatedness of curriculum and assessment, and although the current assessment review and the curriculum review launched on 20 January 2011 are, for obvious reasons, in the wrong order, the published remit for both reviews requires a relationship to be forged between them.

On the basis of the CPR's consultations and activities over the past year or so, we would like to note some additional concerns.

Mindsets and myths

There is a serious and persistent problem of a political mindset – we have encountered it across all the main political parties - which equates assessment with testing and testing with accountability. Thus, when the CPR raised evidentially legitimate questions about the methodology and impact of the KS2 tests in 2007 and 2009 it was accused of being against both assessment and accountability. Though there is less stridency about these matters since the 2010 election, reductionism and technical misunderstanding are never far away. The assessment review has an opportunity to assist the process of re-education which is needed if we are to move to an assessment system which balances the summative and formative, properly links assessment to curriculum, and understands that there are more ways that schools can be called to account than on the basis of test results alone.

Another problem with the current assessment discourse is its habitual recourse to myth and overblown claims, especially where the vexed question of standards is concerned. The CPR's final report identified and disposed of some of the most prominent of these myths and claims – testing of itself drives up standards, parents of young children wholeheartedly support testing, standards in the basics and the pursuit of curriculum breadth are incompatible, teachers before 1997 were 'professionally uninformed', England now has the best teachers, best-trained teachers and highest educational standards ever, and so on.⁶ We hope that the assessment review will set its face against this kind of thing. It is ironic that in respect of assessment, an activity which is above all about the marshalling of evidence, evidence is so often absent from the attendant political discourse and commentators resort to the kinds of off-the-cuff judgement that they say teachers should eschew.

Realistic expectations

As has been argued by the CPR and many others, the current system of KS2 summative assessment in limited aspects of a small number of subjects is made to bear far too big a

⁵ Cambridge Primary Review (2010) *After the Election: policy priorities for primary education*, Cambridge: University of Cambridge Faculty of Education

⁶ CPR final report, pp 473-4.

burden of expectations. KS2 summative assessment should aim to be just that: a procedure for reliably and validly assessing pupils' attainment in specific subjects or competencies at an appropriate point towards the end of KS2 for the purposes of informing pupils, teachers and parents. It should not be expected simultaneously to provide the sole measure of a school's performance, let alone the performance of the education system as a whole; nor should it be treated as a proxy for children's performance in subjects other than those assessed.

It follows that work will be needed (a) to strengthen formative assessment, (b) to identify other indicators of school performance, (c) to identify other procedures for monitoring the quality of the system as a whole. To strengthen accountability we need both effective school governance and appropriate patterns and criteria for school inspection. To gauge system-level performance, as the CPR has recommended, the APU model can be revisited, as long as it is combined with other performance data. In other words, systems are already available and they should be properly used. If we are to improve assessment we must at the same time reduce to a more realistic level the expectations placed upon it by refining and giving due weight to inspection, governance and system-level monitoring.

Apropos the problem of the assessment 'burden', we are also concerned that the first item in the assessment review's remit perpetuates the view, challenged in the final CPR report and above, that assessment of itself raises attainment and narrows the gap between high and low attainers. Good teaching raises attainment; assessment only defines or measures it. Assessment may expose the attainment gap but it cannot close it. Closing the gap is not only about good teaching but also – since we know that the attainment gap maps with dispiriting precision onto the gaps in income, health, wellbeing, risk and opportunity – it is about simultaneously tackling inequality across the full range of public policy.

Reinforcing entitlement

It is now widely accepted – and deplored – that KS2 assessment has become the tail that wags the primary curriculum dog (and the fact that the present government's assessment review preceded its national curriculum review unfortunately manages to reinforce this historic error). Central to the curriculum review process, the CPR has urged, should be the reinstatement of a meaningful concept of entitlement. So if children are statutorily entitled to a broad and balanced curriculum (which currently they are and the Secretary of State has recently assured the CPR they should be⁷), then assessment procedures and content must reinforce this rather than, as currently, undermine it. It is important that this review should lead to a system of assessment which contains procedures for individual pupil assessment that do not distort the curriculum, or perpetuate the view that what is assessed is all that matters educationally or all that needs to be taught with due seriousness. In a mature system of assessment which sensitively combines the different approaches now available, it should be possible appropriately to assess children's progress and achievement in all aspects of the curriculum to which they are entitled.

Redefining 'standards'

As the CPR points out, parents and the public may have been conned by the heavy standards rhetoric of the past decade or so into believing that claims about 'standards' relate to the totality of young children's education. They do not. Not only should assessment in limited aspects of English and mathematics no longer be treated as proxies for the rest of the curriculum, or even the totality of those two subjects, but standards should be made congruent with entitlement. If children are entitled to x, y and z, then they, their parents and their teachers need to know how they are progressing in each of these, and school inspection should in its turn attend to them. Otherwise, there is no way of knowing if children are getting what they are entitled to. And when we come to the standards of schooling (as opposed to the standards of pupil attainment in subjects x, y and z) there is much more to

⁷ Letter from the Secretary of State to the Director of the Cambridge Primary Review, 14 January 2011.

this than can be encompassed by any system of pupil assessment. The reform of assessment must go hand in hand with the reform of curriculum, school governance and inspection.

Strengthening accountability

Too often, as we and others have noted, accountability and testing are equated. We need a broader concept of accountability and a broader range of mechanisms for assuring it. Further, reducing accountability to KS2 assessment in effect makes schools accountable for pupil progress and attainment regardless of factors outside their control which also exert an influence. We repeat: in England - one of the most unequal among the world's richest nations - schools can do a great deal to combat disadvantage but they cannot tackle it on their own, and their impact will be limited if wider social and economic policies run counter to their efforts.

We also find it helpful to distinguish four levels of accountability:

The individual pupil's learning. Teachers need to know how each pupil is progressing and habitually to use information from formative assessment to optimise learning in respect of school and classroom goals. This requires that formative assessment should be firmly embedded in everyday pedagogy, not only through children's written work but also - and perhaps especially - through classroom interaction of a kind that reveals and probes their understandings and misunderstandings.⁸

The individual pupil's achievement. Teachers and parents need to know at regular intervals where each pupil has reached in respect of school and national goals. This requires a valid, reliable and comprehensive system of summative assessment based, as appropriate, on moderated teachers' judgments, tests and other procedures, in relation to whatever are held to be appropriate standards in all aspects of the curriculum to which pupils are statutorily entitled.

The individual school's performance. Parents and governors need to know how good their school is. This requires regular school inspections which focus on the quality of provision across the board - teaching, learning, curriculum, assessment, leadership and so on - and which respects each school's unique circumstances and ethos while enabling parents and others to make valid school-school comparisons.

The performance of the system as a whole. Government needs to know, and the electorate has a right to ask, how the system as a whole is performing. This requires regular national surveys of pupils' attainment in all aspects of their entitlement curriculum, based on sampling of both pupils and assessment items. It also benefits from international surveys of educational achievement, provided that these are used with intelligence and discrimination. However, the electorate also needs to know how well government and policy are performing and what impact they have on the capacity of schools to achieve high standards. System-level accountability is about policy as well as practice, and about the performance of government, national agencies and local authorities as well as schools.

We hope to be able to say more about each of these at the oral session.

There are alternatives

Immediately after the CPR's final report was published we were repeatedly told that 'there is no alternative' to the current assessment regime. Thankfully, the commissioning of this assessment review signals otherwise, and we hope that the panel will produce a statement which offers clear pointers for reform at each of the four levels above.

⁸ CPR final report, chapter 15.

However, there are problems with some aspects of the assessment review's remit and thus with the degree of variation from current practice which may be deemed possible or acceptable. We hope that the review panel will comment on these problems in their report rather than take them as givens. For instance (our comments appear in italics after each element in the remit):

- 'How best to ensure that the system of assessment in primary schools can improve standards of attainment and progress of pupils, and help narrow gaps.' *It is better to concentrate on how assessment can validly, reliably and comprehensively provide information about pupil attainment and progress, acknowledging that it is teaching that makes the difference, and that the attainment gap has social and economic as well as pedagogical causes. Separate efforts are needed to improve pedagogy and address disadvantage.*
- 'How best to ensure that schools are properly and fairly accountable to pupils, parents and the taxpayer for the achievement and progress of every child, on the basis of objective and accurate assessments; and that this reflects the true performance of the school.' *This, as noted above, appears to make assessment the sole basis for school accountability. It also conflates the different levels of assessment we have referred to. The assessment review's remit does not include governance or Ofsted inspection, but it should refer to these.*
- 'How to avoid, as far as possible, the risk of perverse incentives, over-rehearsal and reduced focus on productive learning.' *An appropriate and essential requirement, but it will not be met if assessment continues to ignore a large part of the curriculum that children experience and to which they are entitled.*
- 'How to ensure that parents have good quality information on the progress of their children and the success of schools.' *Also appropriate and essential, though the success of schools must be judged on more than a narrow spectrum of KS2 assessment alone.*
- 'How to ensure that performance information is used and interpreted appropriately within the accountability system by other agencies, increasing transparency and preserving accountability to parents, pupils and the taxpayer, while avoiding the risk of crude and narrow judgements being made.' *Since it is governments and the media who are most prominent in the misuse of performance information and in making 'crude and narrow judgements' about matters educational, recommendations on this will require courage as well as skill. Quis custodiet ...*
- 'How to ensure that tests are rigorous, and as valid and reliable as possible, within an overall system of assessment (including teacher assessment) which provides the best possible picture of every child's progress.' *This of course is the nub.*
- 'How best to ensure that the assessment system allows us to make comparisons with education systems internationally.' *Here there is another risk of conflation; and a further risk of over-interpretation. The international surveys of educational achievement in which England participates provide a useful basis for such comparison, but they are a long way from telling the whole story, and rather too much is allowed to rest on them. Explanations of relative success are usually highly selective and tend to ignore cultural and demographic factors (for example, the fact that many of the top performing countries are small, rich and culturally less heterogeneous than Britain, or that some of them have a decidedly undemocratic polity⁹).*
- 'How to make administration of the system as simple and cost-effective as possible, with minimal bureaucracy.' *Of course.*

20 January 2011

The authors

Robin Alexander is Fellow of Wolfson College at the University of Cambridge, Professor of Education Emeritus at the University of Warwick and Director of the Cambridge Primary Review.

Wynne Harlen is Visiting Professor of Education at the University of Bristol, former Director of the Scottish Council for Research in Education, and research consultant to the Cambridge Primary Review.

Alison Peacock is Head Teacher of The Wroxham Primary School, Potters Bar, and Leader of the Cambridge Primary Review National Network.

⁹ Alexander, R.J. (2010) ' "World class schools" – noble aspiration or globalised hokum?' *Compare: a journal of comparative and international education*, 40 (6), 801-818.