Rethinking Assessment

University of BRISTOL
Graduate School of Education

Discussion paper 6:

National standards and technology enhanced assessment

Q: Can technology enhanced assessment (TEA) be used to improve the quality of national assessment data?

Q: Can technology enhanced assessment provide better diagnostic information about an education system?

This paper considers the following:

- International comparisons
- National standards
- Advantages and disadvantages of TEA
- How can TEA influence national tests?
- A way forward

Key recommendations

- Governments should be encouraged to address the tension between assessment for learning and meeting accountability demands.
- Efforts should be made to explore how TEA can facilitate the measurement of wider outcomes than high-stakes summative assessments.
- Building on the examples of innovative practice that now exist internationally, governments should be encouraged to explore the scope for TEA to facilitate computer-adaptive national testing that enables students to demonstrate their best performance.
- Governments need to invest in improving the data literacy of educators, so that they understand how to use macro-level assessment information effectively.

The monitoring of national standards is now a key and growing component of many education systems. Governments face increasing pressure to assess student performance at every level of the education system.[1] Such assessment data is used as a key indicator of international educational competitiveness

It is argued that the public availability of national standards data increases the transparency and accountability of educational institutions, as well as exerting pressure for their improvement. As yet, however, the use of technology enhanced assessment (TEA) for this purpose has been limited despite early signs that its use could significantly improve both the value and the efficiency of such monitoring.

International comparisons

National educational goals and policies are influenced by the presumption that education contributes to economic and social development. For this reason, governments around the world are demonstrating an increasingly explicit concern that their students should perform well in the international tests of achievement now being conducted by bodies such as the OECD.[2] Such comparisons generate pride or panic in countries that fare well or poorly. They have also resulted in the increasing phenomenon of 'policy-borrowing' in the hope that practices from apparently successful countries can be adopted for the benefit of less successful ones. International comparisons have created high quality data. However, these data typically relate to indicators generated by the international bodies themselves. This is likely to lead to an increasing similarity in educational priorities.[3] In some cases, countries rely on regional agencies such as the Southern and Eastern African Consortium for Monitoring Educational Quality to run their national assessment activities, rather than developing ones that are local to their context.

National standards

National assessments can provide information about the overall effectiveness of the education system in terms of, for example, value for money and parent satisfaction. Such data are also typically used to monitor the effectiveness of individual schools. In some countries such national assessment data is used as the basis of 'league tables' to encourage competition through the public availability of summative assessment data. In other countries, the primary goal of national monitoring is to provide diagnostic data on individual student performance. Some European countries, for example, use a more formative approach to help teachers to focus more effectively on their students' learning needs. Other countries, such as Finland, use national tests to understand the impact of contextual factors on students' learning.[4]

Many criticisms have been made of the effectiveness and Impact of national assessment tests that are so-called 'high-stakes'. High-stakes tests have been seen to narrow the curriculum, limit the use of assessment for learning and demotivate lower-achieving students. Typically, they encourage teachers to 'teach to the test' and students to 'trade for grades'.[5]

Advantages and disadvantages of TEA

Digital technologies have the potential to improve the design, delivery and targeting of national monitoring. Technology enhanced assessment can be used for delivering tests on line and increasingly, for providing adaptive test questions that match the performance of students as they undertake a particular assessment. In the Netherlands, for example,different digital editions of national tests supply extra questions in certain areas that are intended to support students with learning difficulties.[4] TEA can also provide greater efficiency in the administration and marking of tests, the use of new item formats and improved reporting mechanisms.[6] TEA has the potential to offer both more accurate assessment information and provide immediate feedback to the test-taker.

However, the use of TEA for national assessment is also vulnerable to the pitfalls common to technology-enhanced learning which include the cost and maintenance of infrastructure; the availability of technical support and the need for training of staff involved. There are also issues of fairness since students with greater access to technology may be perceived to have an advantage.[7]

How can TEA influence the content and format of national tests?

The use of TEA can help to ensure that National Assessments reflect agreed educational priorities. Recent international trends in national curricula have given increased prominence to a wider range of competences, skills and dispositions as well as subject-based knowledge and understanding. So far these newer curriculum areas have tended to be under-represented in national testing which is often tied to subject-based curricula and learning outcomes that are amenable to measurement through written tests. Of the European Union's list of 8 ' Key Competencies for Lifelong Learning', for example, only three map to specific subjects- the subjects most commonly assessed in national tests.[4] While the ideological debate on what competencies or dispositions really matter continues, it is clear that assessing different kinds of learning outcome on a national basis is both uncommon and poorly understood. New assessment tools that can assess higher-order, more complex thinking are needed but such qualities are difficult to calibrate, measure and evaluate.[8] TEA can help to address this by developing psychometric models that can evaluate competences and use immersive learning environments to elicit and measure such data. Simulations and electronic games can provide opportunities to collect such assessment data.[9]

Scalability is another key issue in the use of TEA for national assessment tests. The e-scape project at Goldsmith's, University of London (gold.ac.uk/teru/projectinfo) for example, has developed an assessment model of creativity and collaboration that uses adaptive comparative judgment to assess online portfolios. The project investigated how possible it would be to use this model at a national level and found positive responses both from teachers judging the portfolios and from awarding bodies looking to integrate the e-scape technologies.[10]

A way forward in the use of TEA for national assessments

Better data collection and analysis, more relevant assessment content and improved testing formats are all possibilities that TEA can offer to improve national monitoring. How effectively these innovations might also enhance teaching and learning is more dependent on the driving forces behind educational policy that may emphasise other priorities than on the possibilities offered by new technologies. Many of these forces – the pressure for modernisation; concern over economic competitiveness; the pressure of international comparisons – are external to the educational systems themselves yet play a significant role in a nation-state's ability to construct its own educational policy and values.

At the present time, national tests provide governments with useful information about the general levels of pupil performance in a particular country. They are currently much less widely used for improving learning. Despite an increasing recognition of the importance of assessment for learning,[11] the growing reliance on standardised international measures works against the creation of national assessment systems that focus on learners and learning.[12] The use of TEA offers the prospect of monitoring practices that are more individually-focused in terms of the collection, analysis and use of monitoring data which may lead to such data being of greater benefit for teachers and learners.

The use of TEA offers the prospect of monitoring practices that are more individually-focused being of greater benefit for teachers and learners

The growing reliance on standardised international measures works against the creation of national assessment systems that focus on students and learning

1 Benavot, A., and Tanner, E.(2007) 'The Growth of National Learning Assessments in the World, 1995–2006.' Background paper prepared for the Education for All Global Monitoring Report, 2008, UNESCO, Paris.

2 See, for example the OECD's annual 'Education at a Glance' publication and large-scale international comparative assessments such a s the 'Programme for International Student Assessment (PISA)'

3 Ioannidou, A. (2007) 'A Comparative Analysis of New Governance Instruments in the Transnational Educational Space: a shift to knowledge-based instruments?' *European Educational Research Journal*, 6 (4), pp 336-347. dx.doi.org/10.2304/eerj.2007.6.4.336

4 Eurydice (2009) *National Testing of Pupils in Europe: Objectives, Organisation and Use of Results*. A report for the Education, Audiovisual and Culture Executive Agency.

5 Polesel, J., Dulfer, N., and Turnbull, M. (2012) *The Experience of Education: The impacts of high stakes testing on school students and their families: Literature Review*. A report for the Foundation for Young Australians and Melbourne Graduate School of Education. whitlam.org

6 Morris, A. (2011) *Student standardised testing: current practices in OECD countries and a literature review*. OECD Education Working Paper No. 65. oecd-ilibrary.org/education/student-standardised-testing_5kg3rp9qbnr6-en

7 Marks, A. M., & Cronje, J. C. (2008). Randomised Items in Computer-based Tests: Russian Roulette in Assessment? *Educational Technology & Society*, 11 (4), pp. 41–50.

8 Shute, V. J., Dennen, V., Kim, Y., Donmez, O., & Wang, C. (2010) *21st Century Assessment to Promote 21st Century Learning: The Benefits of Blinking*. A report for Digital Media and Learning network. dmlcentral.net

9 Oldfield A, Broadfoot B, Sutherland R & Timmis S (2012) *Assessment in a Digital Age: A Research Review*, Paper prepared for the STELLAR network of excellence. bristol.ac.uk/education

10 Kimbell, R., Wheeler, T., Stables, K., Shepard, T., Martin, F., Davies, D., Pollitt, A. and Whitehouse, G. (2009) *E-scape portfolio assessment: Phase 3 report*. Technology Education Research Unit, Goldsmiths University. gold.ac.uk/media

11 Gardner, J ed. (2006) *Assessment and Learning*, London Sage (2006)

12 Hutchinson, C. and Hayward, L. (2005) 'The journey so far: assessment for learning in Scotland' *The Curriculum Journal*, 16 (2), pp. 225-248.

Rethinking Assessment

2012/2013 Series of discussion papers

6. National standards and technology enhanced assessment

Case study:
National testing in Denmark

Denmark has international attention for its use of an online testing system launched in 2010. The tests are 'low stakes,' as individual national test results remain confidential. Denmark uses results to both support educational regulation and as a pedagogic tool to improve and plan teaching and learning. Students respond to online questions using a 'computer-adaptive testing' (CAT) system, which adjusts the sequence of questions to accommodate a student's level of proficiency as they take the test. Thus, test items will vary for each individual. CAT is seen as a more efficient type of test because it matches levels of difficulty in the test with the learner's ability, which reveals more about the learner than a test where items are too easy or difficult. Results and reports are automatically calculated and analysed, so teachers can view the results immediately.

From: Wandall, J. (2011) 'National Tests in Denmark – CAT as a Pedagogic Tool: Danish National School Agency report'. Journal of Applied Testing Technology, 12, pp. 1-21.

Assessment is universally recognised as one of the most important – and powerful – elements of an educational experience. It is also seen as one of the hardest to reform. However, there is an increasingly accepted need for rethinking assessment if it is to keep up with current theoretical, cultural and technological developments affecting teaching and learning.

Digital technologies open up new possibilities for more personalised, immediate and engaging assessment experiences. However, the use of digital technologies for assessment (referred to as 'technology-enhanced assessment') has yet to be 'transformative', with current practices either replicating traditional assessment methods or manifesting in pockets of innovation that are not widespread.

How the potential of digital technologies can best support improved assessment practices and preferred educational outcomes is becoming an issue of increasing importance. An acknowledgement of the potential that digital technologies offer should recognise the complexity of the task, the many factors affecting successful educational change, and the significant ethical questions raised by the use of digital technologies in assessment.

This series of discussion papers draw on a substantial review of literature which aimed to identify the different ways in which technology currently impacts on educational assessment practices and how it could contribute to a new vision for assessment. The review of literature is available at:
bristol.ac.uk/education/research/sites/tea

The following discussion papers have been produced in order to highlight key issues and questions identified by the review of literature:

Paper 1: Transforming education through technology enhanced assessment
Paper 2: Integrating the formative and summative through technology enhanced assessment
Paper 3: Exploiting the collaborative potential of technology enhanced assessment in Higher Education
Paper 4: Learning analytics and technology enhanced assessment
Paper 5: Ethical issues in technology enhanced assessment
Paper 6: National standards and technology enhanced assessment

Series authors:
Patricia Broadfoot, Sue Timmis, Sarah Payton, Alison Oldfield, Rosamund Sutherland
For further information contact Sue Timmis at Sue.timmis@bristol.ac.uk

Acknowledgements:
We would like to thank Elizabeth McNess (University of Bristol) and Palle Damkjær Rasmussen and Christian Ydesen (Aarlborg University Denmark) for their contribution to this paper.

Graduate School of Education
University of Bristol
Helen Wodehouse Building
35 Berkeley Square
Bristol BS8 1JA
Tel: +44 (0)117 928 9000

bristol.ac.uk/education

Supported by:
PEARSON
STELLAR
SEVENTH FRAMEWORK PROGRAMME

DESIGN: CW-DESIGN.CO.UK