

Select Committee inquiry on the purpose and quality of education: NFER response

25th January 2016

Contact: b.durbin@nfer.ac.uk

Introduction to NFER

1. NFER is a charity and the leading independent provider of education research and assessments in the UK. We provide evidence that improves education, learning and the lives of learners. Our insights are relevant and accessible and inform policy and practice. Successive UK governments have used our evidence to inform policy thinking.
2. Through expert research and extensive knowledge of education and assessment, we offer a unique perspective on today's and tomorrow's educational challenges. We draw on trusted relationships, working with a range of influential organisations from government departments to employers; from school leaders and teachers to parents. The breadth of our work enables us to have a systemic view of the education system, linking together evidence from different areas to give a wide perspective.

Overview of our response

3. We welcome the Select Committee's decision to hold an inquiry on the purposes and quality of education. Effective policy must start from a clear set of objectives and provide mechanisms for testing whether these objectives have been met. This testing is important for holding the system to account and for informing better policymaking in future.
4. We agree therefore that it is important and healthy to debate the purposes of education. However, we do not believe this is a debate it is possible to resolve on objective, evidence-based grounds. Rather, it is a matter of ideological preference that goes beyond the remit of academic discussion, and belongs in the realm of democratic scrutiny. One political party may believe the purpose of education to be X, whereas another believes it to be Y. It is for the electorate to choose between these, and for the party elected to develop policies to deliver the purpose identified.
5. For these reasons, as an independent research charity, our response does not offer a definitive view on the Committee's first question: what the purpose of education should be. Rather, our response focuses on questions 2 and 3: measurement and success. We begin by offering some core principles of measurement within the education sector, before providing an overview of some more specific measures relevant to the types of response we anticipate others will be making to question 1.

6. **Our key recommendation is that if, as a society, we expect our education system to deliver more than purely academic achievement, then we should be investing in higher quality, formal measurement of these other objectives.**

Principles of good measurement

7. There are many ways in which measurement plays a part in our education system, from teachers' day-to-day assessment of individual pupils' understanding, through to the national school accountability system, and international comparison surveys. Measurement can be highly influential on how individuals and the system as a whole behave, and can also be technically challenging. It is therefore important that it is done well.
- a. **What we measure should reflect what we value.** For some, there is far too much testing and other forms of measurement built into our education system. However, in some cases this objection does not relate to measurement per se, but rather to what is being measured. It is clear from the vision statements that many schools adopt that much effort and pride goes into achieving more for their pupils than achieving the exam results most prominent in school performance tables. It is essential that we invest in ways to measure the things society values.
 - b. **Measurement should avoid creating perverse incentives.** Ill-conceived approaches to measurement risk unintended consequences. A well-known example of this is the most commonly used measure of a secondary school's performance: the percentage of its pupils achieving five GCSEs (or equivalent) including English and Maths at grade A*-C. The problem with this measure is that it incentivises schools to narrow their focus on a small group of pupils and subjects close to the C/D borderline, at the expense of other pupils and subjects. This issue has been addressed with the introduction of the new Progress 8 measure, but a similar mistake could be made with the introduction of the new definition of coasting schools (discussed in a recent NFER blog¹ and in our response² to DfE's recent consultation).
 - c. **Measurement should recognise the achievements of different groups of young people.** Given the impact of measurement on behaviour and priorities, it is important that it does not inadvertently leads to bias against any group. Measurement can also be used as a tool for addressing inequality. For example, more explicit measurement of outcomes for disadvantaged pupils (particularly those eligible for free school meals), has formed an important part of recent policy moves to improve outcomes for these groups of pupils – alongside funding through the pupil premium.

¹ <http://thenferblog.org/2016/01/08/coasting-schools-and-the-cliff-edge/>

² http://www.nfer.ac.uk/publications/99953/99953_home.cfm

- d. **Measurement should be technically sound.** When considering a particular measurement approach it is important to consider three main attributes. NFER most often considers these in the context of test development, however the principles are more widely applicable to any form of measurement. First, **validity** refers to ‘the degree to which accumulated evidence and theory support a specific interpretation of test scores for a given use of a test’³ – in other words, is our measurement approach successfully capturing what we intend it to? Second, **reliability** refers to ‘the degree to which test scores for a group of test-takers are consistent over repeated applications of a measurement procedure and hence are inferred to be dependable... the degree to which scores are free of random errors of measurement’. Third, **manageability** considers the cost, practicalities and burden associated with measurement – highly valid and reliable measures could be available, but also be impractical to implement.
- e. **Measurement should be used to drive improvement.** Effective feedback loops, whereby the measurement of outcomes is used to drive improvement, should occur at every level in the system. In the classroom, assessment should be used to provide immediate feedback to teachers and students, in order to facilitate the learning process. The Education Endowment Foundation (EEF) toolkit rates effective pupil feedback as one of the highest impact practices⁴. NFER has previously argued⁵ for a more evidence-informed education system, providing mechanisms by which measurement of outcomes across many classrooms can be used to gain new insights into effective practice. Finally, government policy should be informed by effective measurement. This applies at every stage of the ‘ROAMEF’ cycle described in the HM Treasury’s Green Book⁶, from establishing rationale and objectives, through to the development of potential responses and evaluation.

8. Recommendation: All measurement within our system, including any proposed new testing or accountability measures, should be evaluated against these five criteria.

Outcome measures and success

9. In order to structure our review of outcome measures, we have organised possible responses to the question ‘what is the purpose of education’ into four categories. These could be considered to span two dimensions. Horizontally, this covers more traditional and progressive views of the purpose of education. Vertically, this covers outcomes intrinsic to individual young people and the more extrinsic outworking of these.

³ American Educational Association, American Psychological Association and National Council on Measurement in Education (2014). Standards for Educational and Psychological Testing. Washington, DC: American Educational Research Association.

⁴ <https://educationendowmentfoundation.org.uk/toolkit/toolkit-a-z/feedback/>

⁵ <http://www.nfer.ac.uk/publications/99942>

⁶ ROMEF stands for ‘Rationale, Objectives, Appraisal, Monitoring, Evaluation, Feedback’, and forms the common stages of the policy cycle:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf

Table 1: What is the purpose of education?

A. Academic achievement, including the acquisition of basic numeracy and literacy skills	B. Becoming well-rounded, creative, resilient and fulfilled individuals
C. Preparation for employment and contributing to economic growth	D. Becoming responsible citizens, playing an active role in civic society

A. Academic achievement

10. Measures of academic achievement are highly developed and an established part of our education landscape, lying at the heart of our accountability framework. In England there is a baseline assessment for reception age children, a phonics check for children in year one, teacher assessments at the end of Key Stage 1, statutory tests in reading, writing and mathematics at the end of Key Stage 2 (and proposals for re-sits in Year 7), and a succession of examination performance data for Key Stages 4 and 5 and beyond.

11. Nevertheless, tests can only ever assess a sample of the curriculum and children’s skills, and there remain elements of academic achievement which are harder to measure and which do not feature in the current system. These include speaking and listening (recorded separately as an ‘endorsement’ on GCSE certificates, but not counting towards the final grade), extended writing and mental arithmetic (although this will partly be addressed through proposed ‘times tables tests’). There is a danger that, because they are not formally measured, these skills become neglected.

12. Although GCSE results have been improving over recent years, this does not necessarily mean the system is performing better. Our factsheet on GCSE and A-level standards⁷ explains how international survey data suggests that performance has been remarkably stable over recent years. GCSE pass rates are affected by a wide range of other factors besides the underlying standard of education: changes to the curriculum, changes to how exam questions are written, and changes to the policy on re-sits are all relevant.

⁷ <https://www.nfer.ac.uk/publications/FFEE03/FFEE03.pdf>

13. International surveys provide a useful benchmark against which to compare our performance, however great care is required in their interpretation. As explained in another of our factsheets⁸, common mistakes include: interpreting changes in our ranking over time to imply changes in performance, despite this often being primarily a function of which, and how many, other countries participate in a given year; over-interpreting changes in scores over time, or differences with other countries, when these are within the statistical margins of error; and over-simplistic inferences being drawn from policies adopted in high performing countries, without robust evidence of cause and effect, nor sufficient consideration of wider contextual factors.
14. The new National Reference Test, currently being developed for Ofqual by NFER, will be introduced in March 2017 and will be taken by a random sample of students at around 300 secondary schools. Its purpose will be to provide an anchor for GCSE standards, and a means of robustly tracking standards over time.

B. Becoming well-rounded, creative, resilient and fulfilled individuals

15. To a greater or lesser extent, every school states their purpose as being a holistic one, not only seeking to develop the academic abilities of their students, but equipping them with the social, emotional and practical skills needed for life. However, many of these other outcomes are harder to measure, or at least the approaches to measurement are less well developed or widely adopted.
16. One aspect of the wider role for education is students' engagement in learning and their level of interest in continuing to pursue future study or employment. International surveys, such as the Trends in International Mathematics and Science Study (TIMSS), include measures of engagement in learning. The Government has drawn greater attention in recent times to the role that character and resilience plays in young people's success. In January 2015 a £3.5 million character education grant fund was announced⁹, alongside a £one million grant to the EEF to expand research into the most effective ways character can be taught.¹⁰ This built on a literature review¹¹ on the impact of non-cognitive skills on outcomes for young people. The review concluded that, although there is some evidence of such impact, there are substantial gaps in the evidence base. Furthermore, the starting point of the review was to view these skills as a means to an end, rather than as a purpose of education in its own right that should be tracked and monitored over time.

⁸ <https://www.nfer.ac.uk/publications/FFEE04/FFEE04.pdf>

⁹ <https://www.gov.uk/government/news/character-education-apply-for-2015-grant-funding>

¹⁰ <https://educationendowmentfoundation.org.uk/apply-for-funding/character-and-education-funding-round/>

¹¹ https://educationendowmentfoundation.org.uk/uploads/pdf/Non-cognitive_skills_literature_review_1.pdf

17. Tellus was a national survey run in 2007, 2008 and 2009 to track pupils' wellbeing and views about life.¹² No equivalent national survey, suitable for tracking changes over time at a national scale, has replaced Tellus. If, as a society, we value a wider set of outcomes from our education system, we may want to consider reinstating such a survey, which would offer accountability and inform improvement. NFER's pupil attitude survey provides one systematic approach to monitoring a wider set of outcomes. The survey includes measures of wellbeing¹³, and we have used the dataset in the past to draw attention to specific issues, such as bullying and healthy eating¹⁴. However, it is limited to schools opting in to this paid-for service.
18. In the face of growing concerns about young people's mental health, the lack of a well-established approach to tracking outcomes related to mental health over time is perhaps short-sighted. The upcoming survey of the mental health of children and young people being undertaken by the Office for National Statistics (the first such survey since 2004) is therefore a welcome development.
19. There are many other outcomes under this broad category which could be considered, including personal, social, financial and economic education, and other skills such as creativity, communication, and teamwork. A systematic overview of these is beyond the scope of this response, but could form part of the discussion and follow-up to the Select Committee's inquiry.

C. Preparation for employment

20. Academic qualifications alone are not always accepted by employers as being a suitable measure for identifying young people's preparedness for the world of work.¹⁵ One way to measure preparation for employment and contribution to the economy in the long term is to make use of existing datasets on employment outcomes. By linking these to school- and pupil-level data such as Individual Learner Records and the National Pupil Database, it is possible to make inferences about the effectiveness of the system as a whole, individual schools, and (by using econometric analysis) the returns to specific qualifications.¹⁶ Data suitable for this purpose include:
- HMRC data on earnings and employment
 - The Work and Pensions Longitudinal Study
 - Birth cohort studies
 - The Labour Force Survey
 - Learner participation in and destinations for post-compulsory education.
21. Efforts by DfE to provide more systematic data on student destinations represent a welcome innovation, as would continued improvement of and increases in the quality of this data, in order to provide a robust additional lever for accountability and improvement.

¹² http://www.nfer.ac.uk/publications/TEL01/TEL01_home.cfm

¹³ <http://www.nfer.ac.uk/schools/school-surveys/wellbeing.cfm>

¹⁴ http://www.nfer.ac.uk/publications/ASUR05/ASUR05_home.cfm

¹⁵ <http://www.cbi.org.uk/media-centre/news-articles/2012/05/gcses-not-up-to-the-job-bosses-say/>

¹⁶ See for example

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/387160/RR398A_-_Economic_Value_of_Key_Qualifications.pdf

22. Alternative measures for how effectively our education system is preparing young people for work include employer surveys. For example, the CBI runs an annual survey which tracks employers' views on the skills of their workforce, and their experiences of recruiting people with particular skills.¹⁷ Another method is to measure the skills that employees have, such as through the International Survey of Adult Skills¹⁸ which, in addition to assessing adults' skills, collects background information about participants – including education and qualifications – through a questionnaire.
23. Another available measure is the number of young people who do not succeed in progressing into further education, employment or training. This group, known as NEETs (Not in Education, Employment or Training), is tracked by local authorities and schools, with LAs typically collecting and analysing readily accessible data on student characteristics which are strongly correlated with becoming NEET. Schools use their own records and knowledge of young people to moderate the local authority's 'at risk of NEET' list.¹⁹

D. Contributing to society

24. Some areas of the education system specifically aim to teach young people how to contribute to society. One such area that has been comprehensively reviewed by NFER is citizenship education.²⁰
25. Citizenship has been a statutory National Curriculum subject in secondary schools in England since 2002, although there is currently no official citizenship curriculum for primary schools. The UK Government recommended citizenship education for implementing its anti-extremist Prevent Duty that became a legal requirement in schools on 1 July 2015.
26. As part of NFER's work in leading the Citizenship Education Longitudinal Study (2001-2010)²¹, we designed a survey which included items that measured respondents' citizenship knowledge. From this we were able to conclude that young people have variation and gaps in their civic knowledge, particularly in relation to economic issues and how parliamentary bodies are composed. The research concluded that this raises questions about how those gaps can be plugged for both young adults and young people still at school. It suggested the need to ensure that the acquisition of political, legal and economic knowledge remains at the heart of how citizenship is taught and learned both in and beyond school.²²
27. This survey could provide one possible basis for measuring how well our education system is preparing young people to play an active role in society in future. An alternative would be to measure participation outcomes more directly, using studies such as 'Understanding Society' or the 'Longitudinal Study of Young People in England'.

¹⁷ <http://news.cbi.org.uk/reports/education-and-skills-survey-2015/education-and-skills-survey-2015/>

¹⁸ http://www.nfer.ac.uk/publications/IACZ01/IACZ01_home.cfm

¹⁹ <http://www.nfer.ac.uk/publications/INDS01>

²⁰ <http://www.nfer.ac.uk/research/projects/cels-cit/>

²¹ The world's largest and longest-running study about the impact of citizenship education on young people.

²² http://www.nfer.ac.uk/publications/CIVT01/CIVT01_home.cfm

Conclusion

28. There are a large number of outcomes of our education system, and it is important within our democratic society to continue active debates about which are most important. However, for this discussion to be meaningful we need also to be able to measure how well the system is performing against these objectives.

29. There are various existing and potential new measures. In the first instance, we should make better use of data that already exist to build up a picture of how effectively the system is delivering what we expect it to. There are also examples of new measures with the potential to place greater emphasis and focus on the outcomes we value most, but these should be introduced with care, given the technical challenges and dangers of unnecessary burden or perverse incentives.