Issue 17 January 2014

Research Matters

CAMBRIDGE ASSESSMENT

UNIVERSITY *of* CAMBRIDGE
Local Examinations Syndicate

CAMBRIDGE ASSESSMENT

Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate, a department of the University of Cambridge. Cambridge Assessment is a not-for-profit organisation.

Citation

Articles in this publication should be cited as: Suto, I. and Rushton, N. (2013). Educational provision for less able students of English and Mathematics. *Research Matters: A Cambridge Assessment Publication*, 17, 2–8.

If you would like to comment on any of the articles in this issue, please contact Sylvia Green.
Email:
researchprogrammes@cambridgeassessment.org.uk

The full issue and previous issues are available on our website:
www.cambridgeassessment.org.uk/our-research/all-published-resources/research-matters/

Research Matters : 17

A CAMBRIDGE ASSESSMENT PUBLICATION

Foreword

This edition of *Research Matters* has an unusual character, not obvious from the surface. Key elements of the work were commissioned urgently, as part of the now-superseded English Baccalaureate Certificate (EBC) initiative. Each strand of work is naturally interesting in its own right. But overall, the EBC process represented a more subtle yet fundamental shift in the research process around public qualifications regulated by Ofqual. Research on assessment can be seen as a public good, just as research from Higher Education increasingly is being released openly as a public good. But the EBC process was double-edged. On the one hand, it increased the research intensity of the work of the exam boards, each of whom would be competing for the exclusive EBC contracts. Research base and research integrity was clearly a vital part of the submissions. On the other hand, the bidding process commodified that research – bids could succeed or fail on the basis of the underpinning research. Advantage would be gained on the presentation of exclusive, high-quality work revealed only at the point of competitive judgement around the contracts. This process had implications for research focus, research process, and for collaboration across the sector. Of course, we should not be naïve about the 'race to publish' and the 'advantage from discovery' which exists in all areas of enquiry – but intensive commodification of research may carry considerable downsides. For now, we can celebrate the upside of the EBC process – a lot of high quality work was done with amazing intensity – and we can now all enjoy the benefits of the insights gained.

Tim Oates *Group Director, Assessment Research and Development*

Editorial

Most of the articles in this issue report on research related to the current qualifications reform agenda in the UK. Suto and Rushton compare provision for less able students in England and in four high-performing jurisdictions around the world. Although they caution against direct policy borrowing, they recognise that useful insights can be gained through international comparisons. Rushton explores misconceptions in Mathematics through the analysis of errors in responses to examination questions. At a time when qualifications are being reformed and support resources are being developed, research of this kind can inform materials development and help students to acquire a better understanding of mathematical skills and knowledge. In the third article, Greatorex discusses the longstanding debate about the advantages and disadvantages of contextualisation of Mathematics questions. She outlines the difficulty of defining 'context' and presents a number of taxonomies that have been developed to address the challenge of definition. Her conclusions provide guidance for those engaged in qualifications development and the production of high-quality examination questions.

Elliott takes a methodological look at how international comparability studies can best be conducted. She focuses on the advantages and limitations of making descriptive comparisons with other jurisdictions. This is a useful source of information for those who wish to engage in such activities as it highlights the challenges and provides practical advice which can contribute positively to the debate on educational reform. Crisp continues the theme of international comparability on cultural and societal factors in high-performing jurisdictions. She draws together information from a range of secondary sources to provide insights into factors that influence education while recognising that the success of an education system results from complex interactions of different factors and that cultural and societal contexts must be taken into account. Crisp's thoughtful consideration of common themes resulting from her case study approach provides an informative overview of an increasingly prevalent focus of attention.

The qualifications reform agenda has also raised questions about the structure of qualifications. Benton investigates the effects of tiered examinations on the aspirations of young people. He considers factors such as achievement and students' background characteristics alongside entry tier and aspirations to identify potential links to negative effects on wider educational aspirations. The results from this research call into question some widely held beliefs about the impact of tiered qualifications on student aspirations and motivation.

In this issue a number of articles concentrate on current initiatives and medium term changes to qualifications. Dhawan's study takes us further by exploring how developments in neuroscience might affect education and test development in the long term. He provides an overview of brain mapping techniques and the widening range of applications of neuroscience particularly in the fields of education and learning. Dhawan provides some cautionary notes about the limitations of current knowledge whilst also recognising the huge potential for future research.

Sylvia Green *Director of Research*

Educational provision for less able students of English and Mathematics

Irenka Suto and **Nicky Rushton** Research Division

Introduction

Current plans to reform General Certificate of Secondary Education (GCSEs) in England and Wales include a return to linear assessment, the inclusion of more challenging course content, and an increase in demand at the level of what is widely considered to be a pass (Department for Education, 2013). Although these changes may help to stretch the most academically able 14 to 16 year olds, facilitating their progression to A levels and beyond, it is also important to ensure that secondary education caters for the full ability range. Students who struggle with core academic subjects also have a valuable contribution to make to society and the economy. Their educational achievements should be as significant a national concern as those of their more able peers.

In this article, we compare provision for equivalent students in four of the highest performing jurisdictions around the world: Singapore, New Zealand, Alberta (Canada) and Hong Kong[1]. We also explore existing educational provision for less able 14 to 16 year old students of English and Mathematics in England. Although cultural and societal differences provide good reasons to discourage direct policy-borrowing (Crisp, 2014), international comparisons may nevertheless reveal some useful approaches for consideration.

Summary of provision in four high-performing jurisdictions

The main pathways through secondary education in the four high-performing jurisdictions considered in this article are summarised in Figure 1. England has also been included as a comparator. It can be seen that in all four jurisdictions, primary education ends a year later than in England. The point at which, and the extent to which, students are offered different curricula according to their abilities, both vary considerably.

Provision in Singapore

Singapore is a city-state with a population of approximately 5.3 million people (Ministry of Communications and Information, 2013). The Singaporean Ministry of Education is responsible for running state controlled schools and registering independent schools. It is also responsible for developing the school curriculum. Students attend primary schools from the ages of 7 to 12, and attend secondary schools from the ages of 13 to 16. Only primary school education is compulsory, but secondary education is universally available.

In Singapore, students are streamed[2] for the final two years of primary education and for the whole of secondary education. A streamed approach was adopted to solve problems with high dropout rates due to the non-compulsory nature of secondary education, and to allow each student to progress at an appropriate rate (OECD, 2010). The lowest of the three secondary school streams – the Normal (Technical) stream – is for those students with the lowest scores in their end of primary school tests (approximately the bottom 15 per cent). The stream has a vocational focus with a practical teaching approach that is intended to prepare students for further technical and vocational training (Ismail and Tan, 2005).

The Normal (Technical) stream curriculum is designed to ensure that students are proficient in English, Mathematics and Computer Literacy. Students also study their mother tongue language and one or two non-compulsory subjects such as Science, Art, Food Studies, Mobile Robotics and Retail Operations (Singapore Examinations and Assessment Board, 2013). Curricula for all subjects are available from the Ministry of Education (2013) website. The Ministry is currently in the process of replacing the curricula for all secondary school students.

The aims of the post-2013 Normal (Technical) Mathematics curriculum are:

> *to enable students who are bound for post-secondary vocational education to:*
> - *acquire mathematical concepts and skills for real life, to support learning in other subjects, and to prepare for vocational education;*
> - *develop thinking, reasoning, communication, application and metacognitive skills through a mathematical approach to problem solving; and*
> - *build confidence in using Mathematics and appreciate its value in making informed decisions in real life.*
>
> (Ministry of Education, 2012a).

The syllabus builds upon the content from the Foundation Mathematics syllabus which students have followed in primary school. The content is divided into five strands:

i. three content strands:
   a. Number and Algebra
   b. Geometry and Measurement
   c. Statistics and Probability

ii. a context strand (Real World Context)

iii. a process strand (a list of mathematical processes that can be found in the other four strands).

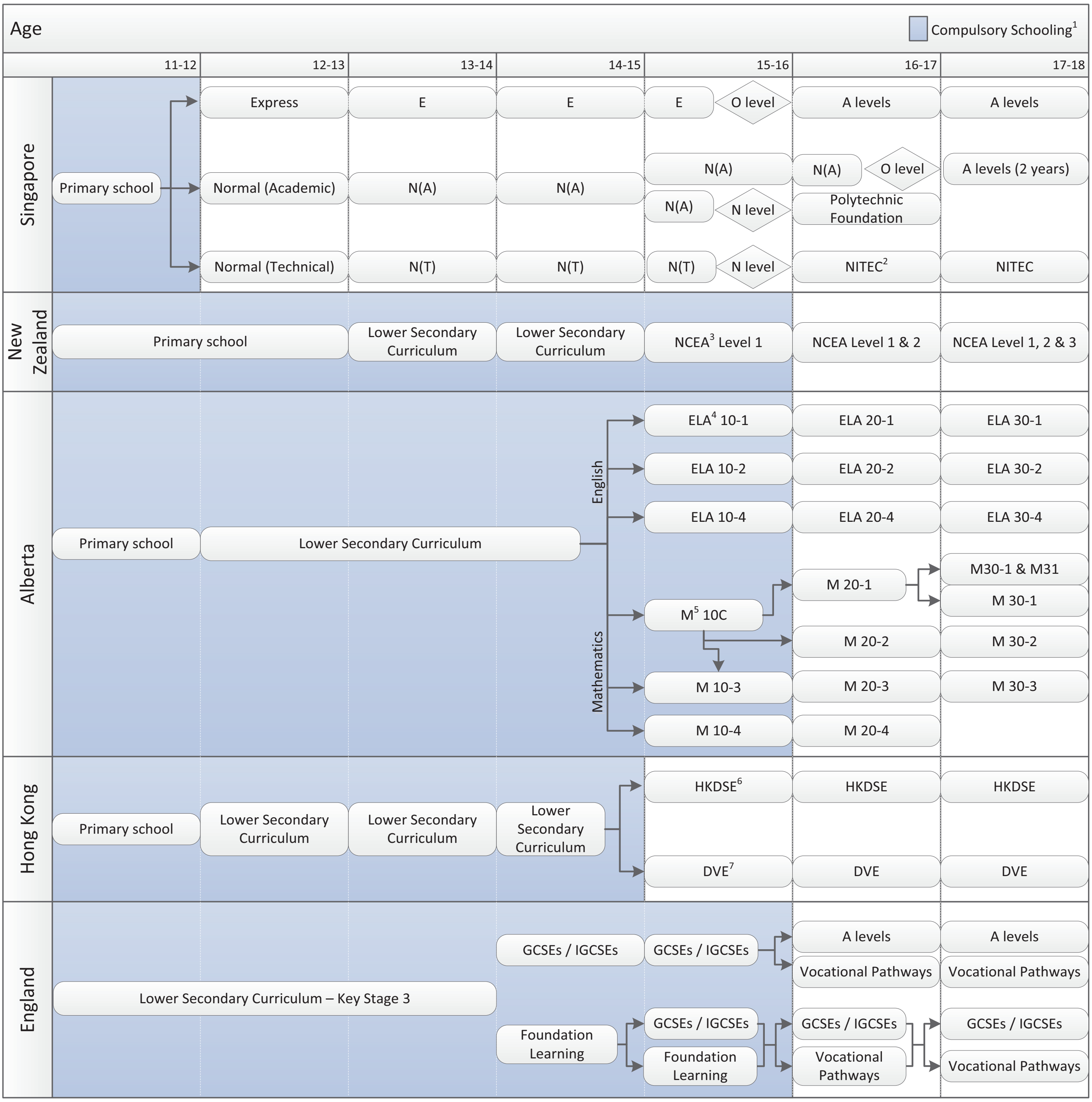
Within the curriculum the content is arranged by strand for each year group, and learning experiences are suggested alongside the content.

All students in the Normal (Technical) stream study English Language. There is no English Literature curriculum for Normal (Technical) stream students, although a few schools offer their own syllabuses as an elective module. There are three aims in the current English Language syllabus, which are intended to result in functional fluency in English:

1. These jurisdictions are considered to be high-performing because they have achieved high scores in recent PISA, TIMSS and PIRLS tests.

2. In this article, streaming is defined as grouping students by ability across all subjects, whereas setting is defined as grouping students by ability for particular subjects.

Figure 1: Pathways through secondary education in four high-performing jurisdictions



1 Note that for England only, compulsory schooling is different from compulsory education. Students can leave school at 16, but they must continue some form of compulsory education until they are 17 (this will be raised to 18 from September 2014). Compulsory education can be an apprenticeship, or full-time employment combined with part-time education/training as well as schooling.

2 NITEC = National Institute of Technical Education Certificate

3 NCEA = National Certificate of Educational Achievement

4 ELA = English Language Arts

5 M = Mathematics

6 HKDSE = Hong Kong Diploma of Secondary Education

7 DVE = Diploma in Vocational Education

1. *Listen, read and view*[3] critically and with accuracy and understanding a wide range of literary and informational/functional texts from print and non-print sources
2. *Speak, write and represent*[3] in internationally acceptable English (Standard English) that is grammatical, fluent, mutually intelligible and appropriate for different purposes, audiences, contexts and cultures

3. "The skills of Viewing and Representing are integrated with Listening, Reading, Speaking and Writing to take into account the importance of developing information, media and visual literacy skills in the teaching and learning of English Language". (Ministry of Education, 2010).

3. Understand and use internationally acceptable English (Standard English) grammar and vocabulary accurately and appropriately as well as understand how speakers/writers put words together and use language to communicate meaning. (Ministry of Education, 2010, p.10).

The syllabus is divided up into six areas of language learning:

i. Listening and viewing
ii. Reading and viewing
iii. Speaking and representing
iv. Writing and representing

v. Grammar

vi. Vocabulary.

Each area of learning is divided into focus areas with associated learning outcomes. The skills, strategies, attitudes and behaviours required for each focus area are listed in the syllabus, and the year groups they apply to (from Primary 5 through to Secondary 4) are identified, thereby showing the progression in skills that are acquired. Within all of the focus areas, students study the use of English in a wide range of texts, including spoken and visual texts (e.g. songs, media programmes, and online texts).

At the end of the Secondary 4 Year, Normal (Technical) stream students sit N(T)-Level examinations. Most subjects are jointly examined by the Ministry of Education and Cambridge International Examinations. Students are assessed via written and practical examinations. In Languages (both English and Mother Tongue) students are also assessed by oral examinations and listening comprehensions. There are five grades: grades A to D (considered to be a pass) and Ungraded. The Normal (Technical) Level Certificate is awarded to all candidates who achieve a pass in one or more subjects. In 2012 it was awarded to 98.1 per cent of the Normal (Technical) students (Ministry of Education, 2012b).

Provision in New Zealand

New Zealand is an island country with a population of approximately 4.5 million people (Statistics New Zealand Tatauranga Aotearoa, 2013). The Ministry of Education oversees the whole education system, and develops the curriculum and the national assessment standards. The New Zealand Qualifications Authority (NZQA) is responsible for: managing the New Zealand Qualifications Framework, administering the secondary school assessment system, recognising qualifications and setting the standards for some unit standards.

Education in New Zealand is compulsory from the ages of 6 to 16. Primary education runs from the ages of 5 to 13 (Years 1 to 8) and secondary education runs from 13 to 18 (Years 9 to 13). There are many different types of schools in the country. The main differences between them are in the age ranges they cater for, whether they are state or independent schools, and whether or not they area Māori schools. Almost all schools accept students of all abilities. Students tend not to be required to repeat years, or to be streamed (although they may be set for particular subjects within their schools).

The New Zealand National Curriculum is followed by all students in New Zealand from Year 1 to Year 13. The content is shown by subject, and within each subject the content is divided into eight levels. Each level covers several year groups, and each year group is expected to be working at two or more levels. As well as English, and Mathematics and Statistics, students must study the Arts, Health and Physical Education, Science, the Social Sciences, and Technology until the end of Year 10.

From Year 11 to Year 13, students work towards National Certificates of Educational Achievement (NCEAs). These certificates were designed to "recognise and credential the learning success of all students, whatever their traditional academic prowess." (Hipkins, 2013, p.19). The NCEAs are available at three levels: Level 1, Level 2, and Level 3. Students are expected to achieve a Level 1 Certificate in Year 11, Level 2 in Year 12 and Level 3 in Year 13. However, the content is not tied to particular academic years. At each level, students are able to study some content at the level below. Students can also take more than a year to cover the content, which means that low performers are able to take more than one year to complete each level of their NCEAs (New Zealand Qualifications Authority, 2013a).

The NCEA subjects are divided into many individual 'standards' (units), each of which targets a specific skill and is aligned to one of the three levels. Students are able to study a subject at more than one level, for example taking History standards at Level 1 and Level 2 during the same school year. Each standard is worth a certain number of credits, with each credit worth approximately ten hours of study. Generally, courses that students follow are worth 18 to 24 credits, although some courses only contain 12 credits and others contain more than 30. To achieve an NCEA certificate, students have to achieve 80 or more credits, of which at least 60 have to be of that level or above. In addition, for the Level 1 NCEA, students must achieve ten of these credits for Level 1 Numeracy and ten for Level 1 Literacy. The standards containing the literacy and numeracy requirements occur in a range of subjects, not just in English, Mathematics, and Statistics.

There are two types of standards: *achievement* standards and *unit* standards. *Unit* standards are based on competency and are usually graded A (achieved) or N (not achieved). They are all assessed internally through the accumulation of evidence. *Achievement* standards are based on the New Zealand curriculum and are graded A (achieved), M (merit), E (excellence) or N (not achieved). Those with content that cannot be tested in an exam (e.g. research projects or speaking) are assessed internally; other achievement standards are assessed externally, usually at the end of the school year.

Students can also take National Certificates (New Zealand Qualifications Authority, 2013b). These qualifications are available in a range of school-related areas and are intended to prepare students for further learning or for a related line of employment. Many of the National Certificates are available in subjects related to particular professions (e.g. the National Certificate in Hospitality, the National Certificate in Building, Construction, and Allied Trades Skills) and most of these are only available at Levels 2 and 3; however, there are also National Certificates in Mathematics available at Levels 1 and 2 (*ibid*). No National Certificate exists specifically for English or Literacy, but literacy skills are covered within some of the National Certificates (e.g. the National Certificate in Employment Skills).

Provision in Alberta, Canada

The province of Alberta in Western Canada is approximately 2.5 times the size of the UK but has a population of only 4 million people (Alberta Government, 2013). Typically, students in Alberta attend a junior high school from Year 7 until the end of Year 9. At the end of Year 9, they sit provincial achievement tests in 'English Language Arts', Mathematics, Science and Social Studies. The tests provide information about students' achievements and facilitate comparisons across the province. Teachers use the test results to reflect on and improve their teaching, as well as to report levels of achievement to students and parents (Alberta Education, 2013a).

Subsequently, most students transfer to a senior high school for Years 10, 11 and 12. Of these final three school years, Year 11 is often considered the most stable and productive year for students. This is because in Year 10 they are finding their feet, and in Year 12 they have the stress and excitement of graduating (attending various festivities), and of applying to university and for other educational or employment opportunities (Alberta Learning, 2003a).

At senior high schools, many courses have a '10–20–30' structure. This means that students typically complete the '10' course in Year 10, the '20' course in Year 11, and the '30' course in Year 12. For particular subjects, including English Language Arts and Mathematics, alternative courses with differing content and difficulty are available to students. The course that a student follows will depend upon his or her career and educational aspirations, achievements in the Year 9 provincial achievement tests, and teacher advice (Alberta Learning, 2003b). The start of Year 10 is therefore a key point of divergence into different educational pathways.

The Albertan education system does not aim to get all students to the same point of learning by the time they leave high school. Instead, a key feature of the system is its 'knowledge and employability' courses, which are available in core subjects (including English and Mathematics) from Year 8 to Year 12 inclusive. These courses offer an important educational pathway for less academically able students, but are not designed for students with special education needs. They are:

> *… intended to provide students with opportunities to experience success and become well prepared for employment, further studies, active citizenship and lifelong learning. Knowledge and Employability courses include and promote:*
> - *workplace standards for academic, occupational and employability skills*
> - *practical applications through on- and off-campus experiences and/or community partnerships*
> - *career development skills for exploring careers, assessing career skills and developing a career-focused portfolio*
> - *interpersonal skills to ensure respect, support and cooperation with others at home, in the community and at the workplace.* (Government of Alberta, 2009, p.3).

Enrolment in one or more knowledge and employability courses is determined individually on a course-by-course basis. The decision is based on each student's achievements and goals, and how those goals relate to the philosophy, rationale and intent of the courses. Students are assessed by their teachers. They aim to achieve a Certificate of High School Achievement (Knowledge and Employability), rather than the Alberta High School Diploma which is obtained by the most able students. It is intended that students who achieve the certificate will progress to employment, further training and courses, or other opportunities not requiring post-secondary education *(ibid)*.

The Knowledge and Employability course in 'English Language Arts' is targeted at students who have experienced challenges or difficulty with their skills such that they have a grade level achievement two to three years below their more able peers. The course aims to show students additional strategies for success in English. Course materials tend to have practical applications and are designed to support development of reading comprehension, communication, and other occupational skills, such as creating brief texts (Alberta Learning, 2003b). Similarly, the Knowledge and Employability course in Mathematics is targeted at students who have a grade level achievement in Mathematics two to three years below their peers. It includes topics such as number, shape and space, patterns and relations, and statistics and probability (Alberta Education 2013b).

Provision in Hong Kong

Hong Kong is a specialist administrative region in China with a population of approximately 7.2 million (GovHK, 2013). It is autonomous from China in all areas except defence and foreign affairs. The education system is run by the Bureau of Education. Schooling is compulsory from the ages of 6 to 15. Primary education runs from the ages of 6 to 12. Secondary schooling starts at age 12 and continues until age 18.

Until 2011, the structure of secondary education in Hong Kong was similar to that in some parts of England. Junior secondary education lasted for three years. After a further two years of education, students took the Hong Kong Certificate of Education Examination (approximately equivalent to O levels). Another two years of study led to the Hong Kong Advanced Level Examination (approximately equivalent to A levels). In 2012, the Hong Kong Certificate of Education Examination and the Hong Kong Advanced Level Examination were replaced by a single form of certification, the Hong Kong Diploma of Secondary Education Examinations (HKDSE), which is obtained at the end of secondary education (at age 18). Assessment for the HKDSE is a combination of public examinations and moderated school-based assessments.

Although the three years of senior secondary education (ages 15 to 18) are optional, it is expected that all students will continue their education through this stage (Hong Kong Examinations and Assessment Authority, 2013). There are no academic requirements for entry to senior secondary education, but within individual schools, places may be allocated on the basis of academic performance. Working towards the HKDSE, senior secondary students study four core subjects: Chinese Language, English Language, Mathematics and Liberal Studies. They are also expected to choose two or three elective subjects. These may include science subjects, languages and applied subjects such as Engineering and Production.

Whilst senior secondary students can study all subjects regardless of their ability level, there is some differentiation of subject content for students of differing abilities. In Mathematics, the core (compulsory) curriculum is divided into three strands: Number and Algebra; Measures, Shape and Space; and Data Handling (Curriculum Development Council and Hong Kong Examinations and Assessment Authority, 2007a). Each of these strands contains foundation and non-foundation topics. These are assessed in separate parts of the two HKDSE examination papers that students take at the end of their courses (alongside moderated school-based assessments). The foundation topics are intended to provide all students, including those who are less able, with coherent knowledge of the essential concepts and knowledge within Mathematics. Less able students may study only the foundation topics, or may study the foundation topics plus some of the non-foundation topics. More able students can also choose to study additional content in one of two optional units: Calculus and Statistics or Algebra and Calculus (*ibid*).

The English curriculum is not divided up as formally. There is a compulsory part of the curriculum which lasts three years (Curriculum Development Council and Hong Kong Examinations and Assessment Authority, 2007b). In the second and third of these years, students also study three elective modules, from a choice of eight modules. There is no official reduction in the curriculum content for lower ability students, but teachers are advised that the curriculum can be adapted and reduced for these students (*ibid*).

The curriculum is organised into three strands, which state the reasons for learning English. The 'interpersonal' strand is about interpersonal communication. The 'knowledge' strand allows students to develop and apply knowledge. Finally, the 'experience' strand requires students to respond and express real and imagined experiences. In addition to the three strands, generic skills and values and attitudes are also key components of the English curriculum (*ibid*).

Table 1: Key features of secondary education in four high performing jurisdictions

	England	*Singapore*	*New Zealand*	*Alberta*	*Hong Kong*
All students aim to get to the same level?	No	No	No	No	Yes
% repeating school years[4]	2.2%	5.4%	5.1%	Canada: 8.4%	15.6%
Provision for less able	Either same qualifications with some higher level content omitted or different qualifications	Different courses and qualifications. Low ability students in Normal (Technical) stream [N(T)]	Low ability students move through curriculum at own rate and work on a lower level standards (units)	Different courses and qualifications with different content and focus. Low ability students follow Knowledge and Employability courses	Same qualifications. Low ability students omit some curriculum content
School structure	Varies	Comprehensive, with streaming within schools	Fully comprehensive	Comprehensive	Unknown
Typical age at which separated	Varies by school	12	Unknown	15, or sometimes earlier, e.g. 13	Unknown
Method of separation	Mathematics and English usually set by ability	Streamed across all subjects	Mathematics and English usually set by ability	Mathematics and English usually set by ability	Unknown. Assume setting within subjects
Basis for separation	Varies, depending on when students are set by ability. Key Stage 2 test results, or internal school tests may be used	Primary school leaving exam results	Unknown	Year 9 assessments or teacher evaluation. Weakest students may sometimes be separated after teacher evaluation at the end of Year 7	Unknown
Movement between pathways	No set pathway	Students transfer out of N(T) stream to the Normal (Academic) stream if they meet the criteria	No set pathway. Students study individual standards at appropriate pace	Students can move between adjacent pathways but this may entail an extra year of study	N/A
English curriculum	Content of GCSE specifications is the same for students of all abilities. Low ability students take foundation tier examinations	Separate curriculum for all three pathways. N(T) students only study English Language	Curriculum split into standards at three levels.	Three main pathways, each with own curriculum	Same curriculum for all students. Teachers adapt curriculum for lower ability students
Mathematics curriculum	GCSE specifications are split into foundation and higher tier content. Low ability students take foundation tier content and examinations	Separate curriculum for all three pathways	Curriculum split into standards at three levels	Four main pathways, each with own curriculum	Compulsory content split into foundation and non-foundation topics. Low ability not taught (all) foundation topics
Qualifications at 16 (or end of secondary school if no intermediate exam)	Most students take GCSEs; low ability students take them at foundation level. Some low ability students take other Level 1 qualifications instead	Exams taken at approximately age 16. High ability take O levels; middle take O level and N(A) level; low ability take N(T) level	NCEA subjects taken at 3 different levels, depending on ability	High and middle ability take Alberta High School Diploma (age 18). Low achievers obtain Certificate of High School Achievement (Knowledge and Employability)	Exams at end of senior secondary only, age 18. All learners take same exam (HKDSE)
Type of assessment	External	External	Combination of external and school-based assessment	Alberta High School Diploma: combination of external and school awarded mark. Certificate of High School Achievement: teacher assessed	Combination of external examinations and moderated school-based assessment

4. Taken from PISA data for the number of students repeating one or more grades (PISA, 2009)

Assessment for the English Language part of the HKDSE comprises public examination papers for:

i reading

ii writing

iii listening and integrated skills

iv speaking.

(Curriculum Development Council and Hong Kong Examinations and Assessment Authority, 2013).

Additionally, for school candidates there is a school-based assessment component which aims to encourage extensive reading and viewing. The reading paper and the listening and integrated skills paper each contain three sections. In each case, the first section is compulsory, then candidates can choose between the second (easiest) and third (hardest) sections. Whilst the second section other assesses only the lowest four grades, the third section assesses the whole range of grades (Curriculum Development Council and Hong Kong Examinations and Assessment Authority, 2007b).

Comparison with provision in England

England has a population of 53.5 million people (Office for National Statistics, 2013), which is considerably greater than the populations of the other jurisdictions considered in this article. In England a wide range of courses and qualifications are available for 14 to 16 year olds. Some less able students are offered extra support in taking mainstream general qualifications in English and Mathematics, such as GCSEs, alongside more able students. Others follow a 'foundation learning' pathway in which they may be offered formal qualifications at lower levels, often focusing on core or functional skills.

Foundation learning is aimed at the weakest 20–25 per cent of learners, including those with special educational needs, those who are at risk of disengagement, and those with 'spiky' attainment profiles (Department for Children, Schools and Families, 2010). Whilst students are significantly more likely to be male than female, gender is not a basis for allocation to a foundation learning pathway. Instead, selection is based upon students having achieved significantly below the national average at Key Stage 3 (at age 14) and/or having particular behavioural characteristics (Allan, *et al*., 2010). Students are usually expected to progress to GCSEs or to a vocational educational pathway such as a traineeship or an apprenticeship, although some may move on to supported employment or independent living pathways.

Foundation learning is intended to support the aims of the National Curriculum for Key Stage 4. For each individual, a personalised programme is developed, which is tailored to his or her particular needs, interests and aspirations. Programmes incorporate three key components:

i. vocational/subject learning;

ii. personal and social development; and

iii. functional skills (in English, Mathematics, and ICT).

The choice of component units and qualifications, including the level and size of those qualifications, should be matched to the student's intended destination.

Although stand-alone functional skills qualifications are available, many non-selective secondary schools opt for their lower ability students to follow GCSE or International (I)GCSE courses in English and Mathematics (which have functional skills integrated within them) as part of their foundation learning pathways. A possible reason for this may be that GCSEs and IGCSEs are perceived to have a greater currency, even at low grades. Some schools enter their least able students for Entry level qualifications instead of, or as well as, GCSEs.

Summary

Further key features of the four high-performing jurisdictions described above are summarised in Table 1, alongside features of the English system. It can be concluded that provision for less able students of English and Mathematics varies considerably both within and across countries. Three of the four high-performing jurisdictions considered in this investigation provide courses specifically for such students, as does England. In Alberta and Singapore, courses for the least able focus on knowledge and skills which will have a practical value in the workplace, rather than on preparing students for higher level academic study in English and Mathematics.

References

Alberta Education (2013a) *Ministry of Education website*. Available online at: http://www.education.alberta.ca/ (Accessed 22 October 2013).

Alberta Education (2013b) *Math: It's everywhere*. Available online at: http://www.education.alberta.ca/teachers/program/math.aspx (Accessed 22 October 2013).

Alberta Government (2013) *Alberta's population reaches four million*. Available online at: http://alberta.ca/release.cfm?xID=350618FCB7CB9-D315-C7B6-3A1E8521B366D28A (Accessed 21 November 2013).

Alberta Learning (2003a) *English Language Arts: Guide to Implementation*. Available online at: https://education.alberta.ca/media/883612/srhela.pdf (Accessed 5 November 2013).

Alberta Learning (2003b) *English Language Arts*. Available online at: http://www.education.alberta.ca/media/645805/srhelapofs.pdf (Accessed 18 October 2013).

Allan, T., *et al*. (2010) *Foundation Learning National Evaluation*. Research Report DFE-RR130. Available online at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/183339/DFE-RR130.pdf (Accessed 26 November 2013).

Crisp, V. (2014) Cultural and societal factors in high-performing jurisdictions. *Research Matters: A Cambridge Assessment Publication*. 17, 29–42.

Curriculum Development Council and Hong Kong Examinations and Assessment Authority (2007a) *Mathematics Curriculum and Assessment Guide (Secondary 4–6)*. Available online at: http://www.edb.gov.hk/en/curriculum-development/kla/ma/curr/ss-math-2007.html (Accessed 22 October 2013).

Curriculum Development Council and Hong Kong Examinations and Assessment Authority (2007b) *English Language Curriculum and Assessment Guide (Secondary 4–6)*. Available online at: http://www.edb.gov.hk/attachment/en/curriculum-development/kla/eng-edu/eng_lang_final.pdf (Accessed 22 October 2013).

Curriculum Development Council and Hong Kong Examinations and Assessment Authority (2013) *Assessment Framework for 2013 HKDSE: English Language*. Available online at: http://www.hkeaa.edu.hk/DocLibrary/HKDSE/Subject_Information/eng_lang/2013hkdse-elang.pdf (Accessed 22 October 2013).

Department for Children, Schools and Families (2010) *Foundation Learning: A guide*. Crown copyright. Available online at: http://archive.excellencegateway.org.uk/media/Foundation%20Learning%20Tier%20Support%20Programme/A_guide_to_Foundation_Learning_DCSF.pdf (Accessed 26 November 2013).

Department for Education (2013) *GCSE reform from 2015*. Available online at: http://www.education.gov.uk/schools/teachingandlearning/qualifications/gcses/a00221366/gcse-reform (Accessed 18 October 2013).

Government of Alberta (2009) *Knowledge and Employability Courses Handbook: Grades 8–12*. Available online at: http://www.education.alberta.ca/media/524889/ke_handbook.pdf (Accessed 22 October 2013).

GovHK (2013) *Hong Kong – the Facts*. Available online at: http://www.gov.hk/en/about/abouthk/facts.htm (Accessed 21 November 2013).

Hipkins, R. (2013) *NCEA one decade on: Views and experiences from the 2012 NZCER National Survey of Secondary Schools*. Report by the New Zealand Council for Educational Research. Available online at: http://www.nzcer.org.nz/system/files/NCEA%20Decade%20On%20Final_web%20%281%29.pdf (Accessed 18 October 2013).

Hong Kong Examinations and Assessment Authority (2013) *HKDSE*. Available online at: http://www.hkeaa.edu.hk/en/hkdse/ (Accessed 5 November 2013).

Ismail, M. & Tan, A. L. (2005) *Voices from the Normal Technical world – An ethnographic study of low-track students in Singapore*. In *Redesigning pedagogy: research, policy, practice conference*, Singapore, May–June 2005. Singapore: National Institute of Education Digital Repository. Available online at: http://repository.nie.edu.sg/jspui/bitstream/10497/214/1/2005v4.pdf (Accessed 26 November 2013).

Ministry of Communications and Information (2013) *About Singapore*. Available online at: http://app.singapore.sg/about-singapore/sg-facts (Accessed 21 November 2013).

Ministry of Education (2010) *English Language Syllabus 2010 Primary (Foundation) & Secondary (Normal [Technical])*. Available online at: http://www.moe.gov.sg/education/syllabuses/english-language-and-literature/files/english-primary-foundation-secondary-normal-technical.pdf (Accessed 18 October 2013).

Ministry of Education (2012a) *N(T) – Level Mathematics Teaching and Learning Syllabus (2013)*. Available online at: http://www.moe.gov.sg/education/syllabuses/sciences/files/normal-technical-level-maths-2013.pdf (Accessed 22 October 2013)

Ministry of Education (2012b) *Press Releases: Results of the 2012 Singapore-Cambridge General Certificate of Education Normal (Academic) and Normal (Technical) Level Examinations*. Available online at: http://www.moe.gov.sg/media/press/2012/12/results-of-the-2012-singapore-.php (Accessed 4 November 2013).

Ministry of Education (2013). Available online at: http://www.moe.gov.sg/education/syllabuses/ (Accessed 18 October 2013).

New Zealand Qualifications Authority (2013a) *NCEA levels and certificates*. Available online at: http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/understanding-ncea/how-ncea-works/ncea-levels-and-certificates/ (Accessed 5 November 2013).

New Zealand Qualifications Authority (2013b) *School qualifications list*. Available online at: http://www.nzqa.govt.nz/qualifications-standards/qualifications/other-school-qualifications/list/ (Accessed 22 October 2013).

OECD (2010) *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education*. OECD Publishing. Available online at: http://dx.doi.org/10.1787/9789264096660-en (Accessed 4 November 2013).

Office for National Statistics (2013) *Population Estimates for UK, England and Wales, Scotland and Northern Ireland, Mid-2011 and Mid-2012*. Available online at: http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/mid-2011-and-mid-2012/index.html (Accessed 5 November 2013).

PISA (2009) *PISA 2009 results: What makes a school successful? Resources, policies and practices (Volume IV)*. Available online at: http://www.oecd-ilibrary.org/education/pisa-2009-results-what-makes-a-school-successful_9789264091559-en (Accessed 22 October 2013).

Singapore Examinations and Assessment Board (2013) *GCSE N(T) – Level Syllabuses Examined in 2013*. Available online at: http://www.seab.gov.sg/nLevel/syllabus/schoolCandidates/2013_GCE_NT_ES_School.html (Accessed 10 July 2013).

Statistics New Zealand Tatauranga Aotearoa (2013) *National Population Estimates: At 30 June 2013*. Available online at: http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalPopulationEstimates_HOTPAt30Jun13.aspx (Accessed 21 November 2013).

Common errors in Mathematics

Nicky Rushton Research Division

Background

When answering Mathematics questions, students often make errors leading to incorrect answers or the loss of accuracy marks. Many of these errors will be random, occurring through calculation errors or misreading of the question, and will not affect many candidates. However, some errors may be seen in a number of students' scripts. These are sometimes referred to as common errors.

At the end of each examination session, an examiners' report is produced for centres. These reports are intended to enable teachers to better prepare their students for future examinations (OCR, 2012). The precise content of these reports varies depending on the subject and awarding organisation, but all generally contain a commentary on the way that students answered the questions. In Mathematics this commentary may refer to the methods that students used to answer the question, but it often also includes details about the common errors that were made by students.

There is a debate in the literature about the difference between errors and misconceptions. Confrey (1990) defines both errors and misconceptions as resulting from the rules and beliefs that students hold, but suggests that the difference is that misconceptions are attached to particular theoretical positions. However, other researchers, such as Nesher (1987), use the term misconceptions to describe systematic errors without reference to a theoretical position. Further researchers, such as VanLehn (1982, in Confrey, 1990) and Brown and Burton (1978, in Dickson *et al*., 1984), use a further term, 'bug', to describe those errors that arise from wrong steps in a calculation procedure.

There is a body of research literature that identifies misconceptions in students' mathematical understandings. For example, Swan (1990) described two sets of misconceptions held by students: those that affected their calculations using the four operations (addition, subtraction, multiplication and division); and those affecting their interpretation of graphs. Other researchers have investigated misconceptions that occur in algebra (e.g. Brown & Burton, 1978). However, many of these studies were carried out in the 1970s and 1980s, and the misconceptions that were identified then may not be as relevant today. Changes to the content of specifications, or the way that Mathematics is taught may have affected the errors that students make, and there may now be previously unidentified misconceptions as a result.

The aim of this study was to identify common errors that have been made in Mathematics exams. It examined all the common errors that students made, regardless of whether they were systematic errors, bugs or slips, as all types could provide useful information for teachers and examiners. It focussed on the General Certificate of Secondary Education

(GCSE) and the International General Certificate of Secondary Education (IGCSE) so that the results could feed into the redevelopment of GCSE Mathematics.

Methodology

Three Mathematics specifications were used in this study: IGCSE Mathematics (0580), GCSE Mathematics A (J512) and GCSE Mathematics B (J567). It was necessary to select two OCR Mathematics GCSE qualifications because a change in specification for first teaching in 2010 meant that there was not a single qualification that covered the entire period. These two qualifications were selected because J512 was a linear specification, and J567 was the new version of that specification. The linear nature of the specifications was important as it limited the number of papers that needed to be investigated; the modular GCSE specification would have had too many papers to investigate in the time available.

Copies of the examiners' reports and exam papers were obtained for all three qualifications. It had initially been hoped that it would be possible to investigate both the June and the November sessions, but this was not possible due to the number of papers that would have had to be investigated. Instead the June sessions were chosen, as this was the main session for candidates taking the exams in England. Cambridge International Examinations' (CIE) IGCSE Mathematics qualification is offered around the world, and different versions of the papers are produced for different time zones. For this particular study, the examiners' reports and question papers for the time zone that included England were investigated[1], as this enabled a direct comparison between the GCSE and IGCSE qualifications. Details of the papers examined are contained in Appendix 1.

Within each examiner's report, any common errors that candidates made were identified. It should be noted that the examiners may have used the term "common error" in different ways. In some instances it may have referred to errors that were made by a high proportion of the candidates. Alternatively, it may have described a high proportion of the incorrect answers for that question. In this case, if most candidates got the question right then the common error may only have been made by a small proportion of the candidates. Whilst the examiners would know which of these applied, as they had access to the item statistics for all the questions, the reports do not always make explicit the proportion of candidates that were affected.

Each error was coded against a theme, sub-theme and example. For the purposes of this study, comments were coded as common errors either when the examiner had referred to them as such, or when the examiner had commented on errors that were made by a number of candidates. The themes were taken from a list of subject areas in Mathematics that were developed by Mathematics specialists within OCR. Additional themes were added during the analysis where it was not possible to find an appropriate theme. More detailed sub-themes emerged from the data and were used to enable more detailed coding of particular issues. The themes and sub-themes do not necessarily correspond to the grouping of topics in the specifications. The question number, paper and year for each common error were also recorded. The coding was initially carried out separately for each paper, to make it easier to record the detail that was necessary.

Once the coding for the individual papers had been completed, the coding was combined to allow any common errors affecting multiple papers, years and qualifications to be identified. Errors that affected more than one question were identified, and these are summarised in this article. The common errors have been listed alphabetically by theme (in bold) and sub-theme (in italic). The papers that they affected are identified in the summaries so that it is clear whether they occurred on GCSE or IGCSE papers, and whether they affected foundation/core tier candidates or higher/extension tier candidates[2]. For the GCSE papers, it has been noted whether the errors occurred on calculator or non-calculator papers when this was considered important for the way in which candidates answered the questions. This was not coded for the IGCSE papers, as they are all calculator papers.

Algebraic fluency

Manipulating expressions

Candidates made several common errors when manipulating expressions and equations. When expanding brackets, a common error made by both foundation tier GCSE candidates and core tier IGCSE candidates was to only multiply out part of the brackets (e.g. expanding $3(2x-5)$ to give $6x-5$). Another common error made by core tier IGCSE candidates was to forget to change the sign when multiplying out brackets, particularly for the second term in the brackets (e.g. multiplying out the second term in $-4(x-y)$ as $-4y$).

Core tier IGCSE candidates found it difficult to simplify expressions involving indices. Common errors were to use the wrong operation for the integers and/or indices, and not to realise that a term without a power needed to be considered as an index of 1 when simplifying the indices.

Question	Examiners' Report comment
15 Simplify **(a)** $3p \times 5p^3$ **(b)** $24q^2 \div 8q^{-3}$ [CIE, June 2009, Paper 11]	*"Adding instead of multiplying the integers and not adding 1 to 3 for the index were common errors. Similar errors often spoilt part (b), namely subtracting integers an adding, or even attempting to divide, the indices."*

Common errors made by the higher tier GCSE candidates were to multiply indices that should have been added, and to divide indices that should have been subtracted. Some higher tier GCSE candidates also carried out the subtraction of the indices in the wrong order, so that they obtained a positive power rather than a negative one. When the powers occurred both inside and outside a bracket (e.g. $(3x^4y)^2$) candidates found simplifying the expression more difficult, often adding the indices rather than multiplying them together.

The higher tier GCSE candidates found indices particularly problematic when they occurred as part of a fraction (e.g. $\frac{2^{3x+2}}{2^{x+5}}$). Most candidates did not know what to do, and whilst the better candidates realised that they had to subtract the powers, most did so incorrectly. Common errors were

1. CIE papers have been offered in time zones since 2009. In 2009 two versions of Papers 1 and 2 were produced, but Papers 3 and 4 were not split into time zones. England fell into the first time zone, so English students would have sat papers 11, 21, 3 and 4. From 2010 onwards, three different time zones were used, and all papers had time zone variants. Students in England would have sat papers 12, 22, 32 and 42.

2. GCSE Mathematics papers are split into foundation and higher tier; IGCSE Mathematics papers are split into core and extension tier.

to add the numerical terms together, or to find the correct numerical term but not include it in the power (i.e. obtaining answers of 2^{2x+7} or $2^{2x}-3$ from $\frac{2^{3x+2}}{2^{x+5}}$)

Extension tier IGCSE candidates also had problems simplifying expressions that involved fractions and indices. A common error was to subtract numerical terms/values rather than divide by them, and/or to incorrectly simplify powers (e.g. simplifying $\frac{8x^6y^5}{2x^4y}$ as $6x^4y^4$ or $4x^2y^5$). The extension tier IGCSE candidates found simplifying expressions containing fractions and negative cubed roots (e.g. $\left(\frac{27}{x6}\right)^{-\frac{1}{3}}$) particularly problematic. Many candidates could find the cube root of individual numeric and algebraic terms in the equation (e.g. finding the cube root of 27 and/or x^6), but did not know what to do with the negative sign in front of the cube root.

Factorisation also caused problems for candidates. A common error made by GCSE foundation tier candidates was only partially factorising expressions (e.g. only taking out one factor from an expression when there were two). The higher tier GCSE candidates sometimes added or multiplied terms instead of factorising them even when this is not possible (e.g. trying to add the terms in the expression $2a^2+8ab$). Both the higher tier GCSE candidates and the extension tier IGCSE candidates did not always notice when an expression was a difference of two squares, leading to incorrect factorisations. When they did notice this, they often made errors in finding the numerical values (e.g. dividing the numerical value by two rather than finding the square root, placing one value outside the brackets).

Question	Examiners' Report comment
11(a) Factorise (i) $4x+14$ (iii) x^2-16 [OCR , June 2011, Paper 4]	*"Common errors in part (a)(i) were* $4(x+3.5)$ *and* $2(x+7)$ *and in part (a)(iii)* $(x+8)(x-8)$, $x(x-16)$ *or* $(x-4)^2$.*"*

The extension tier IGCSE candidates sometimes used the quadratic formula to find the roots of equations when factorising, but then made errors in their factorisation.

Question	Examiners' Report comment
6(c) Simplify $\frac{x^2-16}{2x^2+7x-4}$ [CIE, June 2012, Paper 42]	"Many of these gave the factors as $(x-\frac{1}{2})(x+4)$ and lost the 2 from the $2x-1$."

A common error made by IGCSE extension tier candidates in expressions set out as a fraction (e.g. $\frac{x^2-16}{2x^2+7x-4}$) was to attempt to cancel terms without factorising first. GCSE higher tier candidates also made errors in these expressions, often cancelling out the terms in brackets and ignoring the powers that accompanied them (e.g. simplifying $\frac{2(x-1)^2}{(x-1)}$ to 2^2).

Rearranging equations and formulae

Core tier IGCSE candidates commonly got the sign wrong when moving terms from one side of the equation to the other by addition or subtraction (e.g. making a positive term negative or vice versa). They also did not apply division and multiplication to the whole equation correctly (e.g. changing $z=2x-y$ into $x=\frac{z-y}{2}$ or $x=2z-y$).

Higher tier GCSE candidates sometimes confused operations when rearranging equations, (e.g. subtracting when they should have divided). Another common error was for them to rearrange equations in the wrong order (e.g. to take a square root of one term only before that term was isolated), or to remove terms from inside a bracket before the bracket had been expanded.

Question	Examiners' Report comment
13(d) Rearrange this formula to make r the subject $S=4\pi r^2$ [OCR, June 2012, Paper 4]	*"The square root was dealt with correctly but often before the* 4π. *Many subtracted the* 4π *instead of dividing and those who did divide often did two separate divisions."*
13(b) Rearrange this formula to make h the subject $A=2\pi r(r+h)$ [OCR, June 2010, Paper 3]	*"The majority incorrectly tried to move* r *or* h *from within the bracket before either expanding or dividing."*

Writing expressions and equations from descriptions

All candidates found it difficult to write down expressions and equations, either from worded descriptions, or from diagrams of shapes. There were many misconceptions about the way that numeric and algebraic terms were combined in equations/expressions. Some weaker candidates did not know that $2y$ is $2\times y$. Other candidates thought that $y+y=y^2$ or $y\times 6y=7y$. Some candidates combined terms that should not be combined (e.g. writing $w+z$ as wz or $15a+32$ as $47a$). Another common error was to confuse an expression with an equation. This could prove problematic if they had written an expression rather than an equation and a later part of the question required them to use the equation that they had written (e.g. rearranging it or solving it).

Substituting into expressions and solving equations

One of the most common issues with solving equations was candidates using trial and error to find the answer, rather than solving the equation algebraically. Some of these candidates obtained correct answers, but mistakes were common.

There were two common errors made by foundation tier GCSE candidates when substituting values into expressions. They sometimes used the wrong operation, which led to them adding values they should have subtracted. Some candidates also seemed confused by the process of substitution, adding together the values of the numerical and algebraic terms rather than multiplying.

Question	Examiners' Report comment
13(a) Work out the value of $3a-4b$ when $a=5.5$ and $b=2$ [OCR, June 2012, Paper 2]	*"Lower scoring candidates tended to add the 3 and 5.5 and also 2 and 4, giving an answer of 2.5."*

IGCSE core tier candidates also made errors when substituting into expressions. Their common errors included not working out the part of the equation in brackets first, and adding a value that occurred after the bracket when it should have been used to multiply. Other common errors made by IGCSE core tier candidates were to ignore negative signs in the numbers that they substituted in, and to carry out calculations in the wrong order (i.e. ignoring rules about the order of operations).

In substituting numbers into expressions, higher tier GCSE candidates found it difficult to substitute negative numbers and fractions into equations containing powers. Common errors included giving negative values for the square of negative numbers (e.g. writing $(-2)^2=-4$), and not knowing how to square fractions (e.g. thinking that $\left(\frac{1}{2}\right)^2=1$ or 2.5).

Solving equations

Higher tier GCSE candidates made errors when there were fractions in the equations that they were solving (e.g $\frac{x}{2} - 3 = 5$). A common error was to divide the whole equation by the denominator rather than to multiply by it. Another error was to only multiply part of the equation by the denominator (e.g. multiplying by the denominator in $\frac{x}{2} - 3 = 5$ to give $x - 3 = 10$).

GCSE higher tier candidates taking the non-calculator paper commonly made errors in their numerical calculations when solving equations. Some candidates confused multiplication and division; others confused positive and negative answers. Some candidates appeared not to know what to do when the result of a division was not an integer (e.g. $2x - 3 + 0$) and commonly gave the numerical value from the equation as their answer instead (e.g. solving $2x - 3 = 0$ to give $x = 3$).

When solving simultaneous equations, the core and extension tier IGCSE candidates and higher tier GCSE candidates commonly failed to consistently add (or subtract) all the terms in the equations. When equations did not appear in the standard format, some extension tier IGCSE candidates unnecessarily rearranged them and made errors in the rearrangement or subsequent solution. Higher tier GCSE candidates also commonly lost marks through poor arithmetic or by choosing the wrong method to solve the simultaneous equations (e.g. adding equations that should have been subtracted or equated).

The higher tier GCSE and extension tier IGCSE papers also required candidates to solve quadratic equations. The higher tier GCSE candidates appeared not to know the methods for solving quadratic equations (particularly the quadratic formula), or could not use them to find the correct answers. The IGCSE extension tier candidates appeared to know the quadratic formula, but when using it either lost marks for accuracy or made mistakes with negative signs.

The higher tier GCSE and extension tier IGCSE candidates also made common errors on the questions with inequalities in them. A common error amongst both sets of candidates was to change the inequality signs either into a different (incorrect) inequality sign, or into an equals sign. Candidates on the extension tier IGCSE did not recognise that an instruction to "solve $9 < 3n + 6 \leq 21$ for integer values of n" meant that they needed to list the values, and therefore they lost marks for leaving the answer as an inequality.

Appropriate use of calculator

A common error made by core tier IGCSE candidates was to incorrectly input calculations into their calculators, which meant that the order of operations was incorrect.

Question	Examiners' Report comment
1 Work out the value of $\frac{48}{19.1 - 3.5 \times 4.6}$ [CIE, June 2012, Paper 12]	*"A common error was 48 ÷ 19.1 – (3.5 × 4.6) to give –13.5. Also seen a number of times was the denominator worked as (19.1 – 3.5) × 4.6."*

Candidates taking the IGCSE extension paper made errors when squaring negative numbers on their calculators. They commonly input the calculation without brackets (i.e. calculating -5^2 instead of $(-5)^2$).

Coordinate geometry and transformations

Coordinate geometry

Candidates on the higher tier GCSE paper made errors with 3D coordinates. A common error was to confuse the order of the coordinates when writing them down. Some candidates wrote the value for the z axis second, possibly confusing it with the vertical y axis in 2D coordinates.

Transformations

Both GCSE and IGCSE candidates on all papers struggled with describing transformations. One of the most common errors was to misread the instructions in the questions and to give combinations of transformations rather than the single transformation that was required by the question.

For line symmetry, candidates on all papers reflected shapes in the wrong line when they were required to draw a reflection, or drew a translation instead. When identifying reflections they found it difficult to give the equation of the line that the shape was reflected in. Common incorrect lines for the reflection were the x axis or the y axis. Similarly, when asked to describe rotational transformations, candidates often gave an incorrect centre of rotation.

The enlargement questions also caused problems for candidates. When describing enlargements, core tier IGCSE candidates often omitted the centre of enlargement from their descriptions. They also found fractional and negative enlargements problematic, confusing the two.

Question	Examiners' Report comment
5(e) Draw the enlargement of shape A, centre (-4, 8), with scale factor $\frac{1}{2}$ [CIE, 2009, Paper 3]	*"Whilst most candidates were able to demonstrate their understanding of enlargement, often this was with scale factors of either 2 or ½ or from a different centre. A significant number omitted this part of the question."*

Extension tier IGCSE candidates were able to draw and describe enlargements. However, they could not calculate enlargements of area or volume correctly, when given the enlargement of a line. They used the linear scale factor, rather than squaring (for area) or cubing (for volume).

The extension tier IGCSE candidates found transformations involving stretch and sheer difficult, and often confused the two transformations. Another error was to identify stretch or sheer as an enlargement instead. Candidates did not always describe the invariant line correctly.

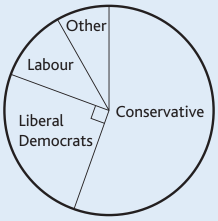
Question	Examiners' Report comment
4(b) Describe fully the single transformation that maps **(ii)** triangle *T* onto triangle *V*. [CIE, June 2010, Paper 42]	*"...an incomplete invariant line e.g. invariant in x rather than the x-axis. Answers such as the x-axis or parallel to the x-axis were not accepted as invariance was not clear."*

Data handling, statistics, probability and chance

Drawing charts and graphs

Candidates on all papers of the GCSE and IGCSE qualifications are expected to draw and interpret charts as part of their data handling skills. The errors that they made depended on the level of the paper, and whether they were GCSE or IGCSE candidates.

Foundation tier GCSE candidates found it difficult to use pie charts to answer questions. Many found it difficult to find the numbers represented in a sector of the pie chart.

Question	Examiners' Report comment
17(c) In one constituency 53 520 people voted in the 2010 general election. This pie chart summarises the results.  **(iii)** How many people voted Conservative? [OCR, June 2012, Paper 2]	*"A common wrong answer was 26 760 from just halving the 53 520. Some did gain credit for the angle of 200°. A common error was to calculate for an angle of 202.5° from 180° + 22.5° rather than using their protractor to measure the angle."*

Higher tier GCSE candidates commonly drew histograms incorrectly. Whilst the width of the bars was usually correct, the heights were usually incorrect because the frequency density was calculated incorrectly. The common error was to treat all classes as though they were of equal width when calculating frequency density (dividing all the frequencies by the same number).

IGCSE candidates and higher tier GCSE candidates made errors when drawing lines of best fit onto scatter graphs. The common errors were joining the points instead of drawing a line of best fit, or drawing a line of best fit that went through the origin. The IGCSE core tier candidates had difficulty describing the correlations that were shown on scatter graphs. Candidates either gave lengthy word descriptions instead of the proper terms (earning them no marks), or gave more detail than was necessary.

Mean, median, mode and range

GCSE and IGCSE candidates at all levels made errors when working with these statistics. Core tier IGCSE candidates commonly confused mean, median and mode. The GCSE foundation tier candidates and IGCSE core tier candidates also had a poor understanding of range, commonly writing down the smallest and largest value, rather than calculating the difference between the two.

All candidates had problems finding the median of data when there was an even number of values. Candidates often gave one of the two middle values as their answer, instead of calculating the mean of the middle two numbers.

Candidates on all papers also found it difficult to find the mean of grouped data. Two common errors were observed in these questions. Candidates often divided incorrectly to calculate the mean, usually either dividing the frequency by the number of classes or using cumulative frequencies. Where the classes represented a range of data, candidates used incorrect mid-points (e.g. the upper bound) to calculate the mean.

Probabilities

Candidates made two types of common error when calculating probabilities. Higher tier GCSE and extension tier IGCSE candidates often appeared not to know whether the sampling was with or without replacement, which led to them using an incorrect denominator in calculating their probabilities. There were further errors made by candidates on all the papers when they calculated combined probabilities. Some candidates used the number of options out of the total possible to give the combined probability, rather than adding/multiplying the probability of each option happening. Other candidates failed to account for all the possible ways in which combined events could occur, leading to incorrect probabilities being calculated. Finally, some candidates got confused about whether the probabilities should be added or multiplied, and used the wrong calculation to find the final probability.

Question	Examiners' Report comment
10(b) Amir tests a laptop at random. Find the probability that both the hard drive and the screen are **not** faulty. [OCR, June 2011, Paper 3]	*"Most candidates identified the correct probabilities part (b) to use in their calculation and those that multiplied them usually reached the correct answer … The main error was to add the probabilities…"*

Ratio

GCSE candidates on both the higher tier and foundation tier papers found it difficult to carry out calculations involving ratios. A common error made by foundation tier candidates was to multiply (instead of divide) the amount by the ratio that they need to find. The higher tier GCSE candidates divided the amount by the required part of the ratio, rather than by the total of the ratio.

Question	Examiners' Report comment
5(a) In a carton of *Squashy*, orange juice and water are mixed in the ratio of 3:7. How many litres of orange juice are needed to make 60 litres of *Squashy*? [OCR, June 2009, Paper 3; also OCR, June 2009, Paper 1 question 17]	*"Weaker candidates tried to divide 60 by 3 and by 7 in part (a)."* (Paper 3) *"The most common working was 60 × 3 = 180, 60 × 7 = 420 giving an answer of 180."* (Paper 1)

Finding correct numerical answers without a calculator

General comments on the GCSE non-calculator papers suggested that candidates' arithmetic skills were poor. The foundation tier candidates had particular problems with subtraction, commonly taking the smaller digit away from the larger one. They also had problems with the order of operations, carrying out calculations in order from left to right, ignoring rules about the order of operations.

Both higher and foundation tier GCSE candidates had problems when decimals were involved in calculations. Candidates on the higher tier could not multiply by decimals correctly, sometimes dividing instead, or on other occasions putting the decimal point in the wrong place.

Fractions, percentages and decimals

Equivalence of fractions, percentages and decimals

Candidates on all the GCSE papers made errors when converting between fractions, decimals and percentages. Candidates often confused tenths and hundredths when converting fractions to decimals (e.g. changing $\frac{3}{50}$

to $\frac{6}{100}$ then writing it as 0.6 rather than 0.06). Higher tier candidates struggled with recurring decimals, often writing degree signs instead of the recurring decimal sign or writing too few digits in the recurring decimal.

Fractions

Candidates on the GCSE non-calculator papers appeared to have several problems when carrying out calculations with fractions without using a calculator. The GCSE foundation tier candidates often failed to put their fractions into simplest form, losing a mark for accuracy. They also did not have an awareness of the relative size of fractions and did not know which numbers to divide and multiply by to find a fraction of a number.

Question	Examiners' Report comment
2(c) Write down a fraction that is smaller than $\frac{1}{10}$ [OCR, June 2009, Paper 1]	*"Roughly half of the candidates not showing an awareness of relative size of a fraction. 0.5/10 was a common answer from weaker candidates."*
11 Work out **(d)** $\frac{5}{6}$ of 78 [OCR, June 2009, Paper 1]	*"Few candidates seemed sure of the method required to do the calculation in part (d). Many were unsure which number, 5 or 6, to divide by (some tried both). Wrong ideas, including 78 – 5 and 78 – 6 were seen"*

The higher tier GCSE candidates found addition and multiplication of fractions without a calculator problematic. Most of the examiners' comments did not identify the exact issues that candidates had, only noting that they were unable to carry out the calculations accurately. However, it was clear that pupils did not have a clear understanding of how to add, subtract, multiply and divide using fractions.

Question	Examiners' Report comment
1(b) $3\frac{1}{3} - 1\frac{5}{6}$ Give your answer in its simplest form [OCR, June 2012, Paper 3]	*"Common errors were doubling $3\frac{1}{3}$ to get $6\frac{2}{6}$, inverting the second fraction and subtracting denominators as well as numerators"*

Percentages

Generally, candidates were able to calculate percentages of numbers, although the GCSE foundation tier candidates were not able to use efficient methods to find them using a calculator.

Question	Examiners' Report comment
12 800 people were asked where they went on their summer holidays. These are the results. UK: 220; Europe: 450; Rest of world: 143; Did not go on holiday: 67 **(a)** Work out the percentage of the people in the survey who replied **(ii)** Rest of the world [OCR, June 2011, Paper 2]	*"Solutions involving long and inefficient methods were often seen in part (a)(ii) and correct answers were rarely seen… Generally candidates did not demonstrate techniques that can be carried out with a calculator in a concise and efficient manner."*

GCSE higher tier candidates and IGCSE core tier candidates found calculating a percentage increase problematic. These candidates were able to find the increased amount, but a common error was to fail to add it onto the total. IGCSE candidates on the extension papers made errors when calculating reverse percentages. Commonly they calculated the percentage of the value they were given instead of using the percentage to find the original value.

Question	Examiners' Report comment
1(b) **(ii)** The money remaining from the \$150 is $37\frac{1}{2}$% of the cost of a day trip to Cairo. Calculate the cost of the trip [CIE, June 2009, Paper 4]	*"However many candidates calculated $37\frac{1}{2}$% of \$30." [\$30 was the money remaining].*

Candidates taking the IGCSE core and extension papers had to solve problems involving simple and compound interest. The common error in these questions was to get the two types of interest confused.

Functions and graphs (linear, quadratic, cubic, trigonometric, exponential)

Functions and graphs

Some of the foundation tier GCSE and core tier IGCSE candidates lost marks when drawing graphs of functions. Common errors occurred when drawing the lines between points. Some candidates omitted lines altogether or drew them freehand. Others joined the dots with straight lines instead of drawing a smooth curve, or drew bumpy curves.

Some core tier IGCSE candidates did not know what the equations for curved and straight line graphs looked like, commonly not recognising the link between the equation that they were given and the graph. When asked to write down the equation of a straight line, candidates could identify where it crossed the axis, but could not give the correct equation. They used the wrong axis in their equations (e.g. $y = -0.5$ instead of $x = -0.5$), omitted the negative sign (e.g. $x = 0.5$ instead of $x = -0.5$), made y a multiple of x (e.g. $y = 0.5x$ instead of $x = -0.5$) or wrote an expression rather than an equation (e.g. $x -0.5$ instead of $x = -0.5$). Higher tier GCSE candidates and extension tier IGCSE candidates made mistakes when calculating gradients of lines. They also commonly wrote fractional gradients as negative when they should have been positive (e.g. $-½x$ instead of $½x$).

Higher tier GCSE candidates and IGCSE candidates also had difficulty with solving equations graphically. The common error made by candidates on both papers was to solve the equations algebraically instead.

Speed

Candidates on the foundation tier GCSE paper found questions about speed difficult. Some candidates who knew the formula for speed and wrote down a distance speed time triangle then used the wrong operation (e.g. multiplication rather than division) to find the answer. Other candidates did not know the formula, so could not answer the question. The extension tier IGCSE candidates did not know how to find distance from a linear, non-horizontal speed-time graph. They incorrectly used the formula (taking values from the graph) rather than finding the area under the graph.

Geometry (including trigonometry and Pythagoras)

Names and properties of shapes, lines and angles

Foundation tier GCSE candidates and core tier IGCSE candidates sometimes made errors in recognising shapes. Commonly confused shapes included trapezium and parallelogram, hexagon and octagon, scalene and isosceles triangles, and isosceles and equilateral triangles. The core tier IGCSE candidates also found it difficult to recognise the labelling of angles, commonly identifying incorrect angles when three letter angle notation was used.

Foundation tier GCSE candidates and core tier IGCSE candidates confused area and perimeter or area and volume. These candidates also made errors when calculating the volumes of prisms. The IGCSE candidates commonly used an incorrect formula in their calculations. Some GCSE candidates were not able to use the formula they had been given to work out the volume, possibly because they did not know how to calculate the cross-sectional area.

When working with circles or cylinders, a common error on all GCSE papers and the core tier IGCSE papers was to confuse diameter and radius. The formulae for circumference and radius of a circle were commonly confused by candidates on all papers.

The properties of circles proved problematic for the core tier IGCSE candidates in questions where they had to find/calculate indicated angles. Candidates often made incorrect assumptions about the lines or shapes that were shown in the diagram (e.g. failing to recognise a line as the diameter or assuming that a triangle was isosceles). Extension tier IGCSE candidates often knew the properties of lines and circles and were able to use them to find angles. Their common error was not using the correct terminology to describe their reasoning (e.g. using terms like z angles instead of alternate angles).

Core tier IGCSE and foundation tier GCSE candidates did not always know the sum of angles in shapes and on straight lines. GCSE foundation tier candidates did not know the angle sum of the interior angles in triangles and quadrilaterals, confusing 180˚ and 360˚. IGCSE core tier candidates commonly confused exterior and interior angles. They also used 180˚ instead of 360˚ in the formula for calculating the size of exterior angles. Some GCSE candidates thought all the angles on a horizontal line added up to 180˚, regardless of whether the angles were around the same point, or related by properties of lines (e.g. whether they were corresponding or alternate angles).

Loci

Questions about loci proved problematic for candidates on all papers. Candidates did not know which construction to use, and commonly drew incorrect ones to solve problems. Finding angle bisectors appeared particularly problematic. In addition, some candidates lost marks because they did not use compasses.

Trigonometry

GCSE higher tier candidates and IGCSE candidates on both tiers sometimes did not know which trigonometric ratio to use to answer questions. Candidates assumed that triangles were right angled when they were not, or failed to recognise right angled triangles in more complex shapes. Another error made by GCSE higher tier candidates was to calculate the cosine of an angle when they needed to find the inverse cosine.

Bearings

Candidates on all papers made both possible common errors when measuring bearings. They failed to measure bearings from north and they measured the bearings in an anticlockwise direction instead of clockwise.

Knowledge of realistic answers

Foundation/core tier candidates at both GCSE and IGCSE often failed to consider whether the answer that they obtained was realistic.

Matrices

Matrices only appeared on the IGCSE extension tier papers, but they led to several common errors. Candidates knew how to find an inverse of a matrix, but made errors rearranging the matrix. Several of the matrix questions involved algebraic expressions, and in these questions candidates sometimes failed to simplify their expressions. Often this occurred because the candidate had given their answer as a 2×2 matrix, when a 2×1 matrix was the correct answer.

Measures

Candidates across all the papers gave incorrect units in their answers. Foundation tier GCSE candidates commonly gave either cm or cm^2 instead of cm^3 as the unit for a volume they had calculated. Another common error was for candidates to misread the type of units that were used in the question and to give the wrong unit in their answer (e.g. giving an answer in cm^2 instead of $feet^2$).

Conversion between units also caused problems across all the papers. IGCSE and GCSE candidates on both tiers commonly forgot to convert their answers to different units after completing their calculations. The GCSE foundation tier candidates made mistakes when converting distances between cm and km, often giving answers in the wrong order of magnitude.

Question	Examiners' Report comment
15(b) Parvinder has a bicycle. Each wheel has a diameter of 65.5cm. On one journey each wheel rotated 3509 times. Calculate the distance Parvinder cycled. Give your answer in kilometres. [OCR, June 2012, Paper 2]	*"Some attempted to convert the units but most were divided by just 1000 or 10 000 rather than 100 000."*

Higher tier GCSE and extension tier IGCSE candidates struggled to convert units when calculating areas or volumes. The common error was using the linear conversion (e.g. 10cm = 0.1m) rather than squaring or cubing these conversions.

All candidates appeared to have problems with questions about the bounds of measurements. Identifying the upper bound was problematic for GCSE higher tier candidates and IGCSE core tier candidates. They often gave incorrect answers such as 12.4 or 12.49 instead of the correct answer of 12.5. (Generally candidates were able to identify the lower bound correctly.) The IGCSE candidates made errors when they used bounds to find the maximum and/or minimum values for calculations.

Common errors were when candidates calculating the bounds after completing the calculation (i.e. giving the maximum and minimum values for the final answer only), or choosing the wrong bounds to find the maximum difference.

There appeared to be a number of errors when candidates were asked questions involving times on a calculator paper. A very common error (made by all IGCSE candidates and higher tier GCSE candidates) was to calculate a time assuming that there were 100 minutes in an hour rather than 60. Working with times over two days caused problems for the core tier IGCSE candidates, who often failed to give the correct day of the week. Generally they were out by one day, or gave the number of days rather than the name of the day.

Question	Examiners' Report comment
3 The ferry from Helsinki to Travemunde leaves Helsinki at 17 30 on a Tuesday. The journey takes 28 hours 45 minutes. Work out the day and time that the ferry arrives in Travemunde. [CIE, June 2012, Paper 12]	*"It was common to see Thursday, as well as the number, 1 or next day, rather than the correct day of the week [Wednesday]."*

Powers and roots

Foundation tier GCSE candidates on both calculator and non-calculator papers commonly doubled numbers instead of squaring them, and multiplied numbers by three rather than cubing them. When calculating the roots of numbers, these candidates often divided the numbers by 2 or 4 instead of calculating the square root.

Fractional and negative powers proved problematic for core tier IGCSE candidates and higher tier GCSE candidates. A common error was to divide by 2 when the square root was written as a fraction (e.g. $16^{\frac{1}{2}}$) rather than calculating the square root of the number. Negative powers were commonly interpreted as giving the same number as positive powers but making the answer negative.

Question	Examiners' Report comment
18(a) Evaluate. $64^{\frac{1}{2}} \times 2^{-4}$ [OCR, June 2010, Paper 3]	*"$64^{\frac{1}{2}}$ was often given as 32 and 2^{-4} became -8, -16 or 0.0002."*

Some IGCSE candidates on the extension paper did not know how to calculate using indices when the answer had to be left in the same form. The common error was to think that the rule of multiplying the bases and adding the indices applied when the numbers had different bases.

Question	Examiners' Report comment
5 Write $2^8 \times 8^2 \times 4^{-2}$ in the form 2^n. [CIE, June 2010, Paper 22]	*"…about a quarter of the candidates multiplied the bases together and then added the indices."*

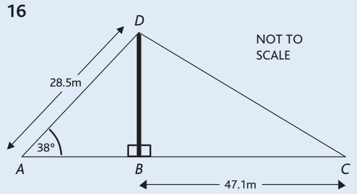
GCSE higher tier candidates did not know how to multiply powers over brackets, commonly adding the indices (e.g. writing the answer for $(7^3)^5$ as 7^8).

In general the examiners' comments suggested that GCSE higher tier candidates were not familiar with the rules concerning surds. Common errors included leaving a square root in the answer when it could have been solved to find a whole number (e.g. writing $\sqrt{9}$ instead of 3), and not simplifying surds (e.g. leaving an answer as $\sqrt{18}$ instead of $3\sqrt{2}$).

Rounding

GCSE and IGCSE candidates made several common errors when required to round their answers. The most frequent error, made by GCSE and IGCSE candidates on both tiers, was failing to give calculator answers to an appropriate number of decimal places (or significant figures). A common error made by the IGCSE candidates was to round incorrectly, or to round too early in their calculations.

16



Question	Examiners' Report comment
16 A flagpole, BD, is attached to level horizontal ground by ropes, AD and CD. AD = 28.5m, BC = 47.1 m and angle DAB = 38°. Calculate **(a)** BD, the height of the flagpole [CIE, June 2011, Paper 12]	*"Unfortunately some lost a mark by presumably rounding 17.546 to 17.55 and then 17.6 to 3 significant figures."*

Finally, questions on the IGCSE extension papers that required candidates to show that answers were equivalent to a rounded value caused problems for candidates. Often candidates failed to show a more accurate value before giving the rounded value, which meant they lost the accuracy mark.

Question	Examiners' Report comment
6 A spherical ball has a radius of 2.4 cm. **(a)** Show that the volume of a ball is 57.9 cm^3, correct to 3 significant figures. [CIE, June 2010, Paper 42]	*"The accuracy mark was often lost as the majority wrote down the answer of 57.9 given in the question. It is important for candidates to understand that if they are required to show how to obtain an answer that is given in the question and is not exact, then a more accurate answer is needed for full marks."*

Sequences and patterns

Candidates at both GCSE and IGCSE found it difficult to give term-to-term rules and the nth term rule for sequences. A common error was to confuse the two. When asked to give the nth term rule, candidates often incorrectly gave the term-to-term rule algebraically instead. Another common error was to give the terms in the nth term rule in the wrong order (e.g. giving 21-4n instead of 4n-21).

Sets

Extension tier IGCSE candidates confused the union and intersect symbols and failed to shade the correct regions of Venn Diagrams.

Standard form

Both GCSE and IGCSE candidates at all tiers found it difficult to write and calculate with numbers in standard form. Answers were commonly written with only the number and power (e.g. 3^8 instead of 3×10^8) or with the decimal point in the wrong place (61.4×10^5). Another common error was to round to the required order of magnitude but to forget to write it in standard form (e.g. 61,400,000 instead of 6.14×10^8).

Terminology

Candidates on all the GCSE papers and candidates taking the core tier IGCSE papers appeared to be confused about factors and multiples, and it was common for candidates to give the wrong one. The highest common factor and lowest common multiple were also sometimes confused, with candidates giving the lowest common factor instead of the lowest common multiple.

Vectors

Core tier IGCSE candidates found vector notation problematic. When using vectors as part of translations, common errors were to misread the scale, to get negative and positive directions confused and to present the translation as coordinates instead of a column vector. Higher tier GCSE candidates using vector notation confused the position and/or the signs of the two values.

The IGCSE extension tier candidates did not recognise some of the properties of vectors. They did not recognise the modulus of vector notation, and they did not realise that a position vector is given from the origin. The comments from examiners suggested that candidates needed encouragement to show their route when answering questions about vectors.

Conclusion

This article has identified the common errors that candidates have made in many different areas of Mathematics. The information may be used to inform future teaching, as the errors can be used to ensure that candidates have a better understanding of the mathematical skills and knowledge that they will need in examinations. It is also useful for awarding organisations as they can use it in the development of future assessment material to ensure that the questions address areas of students' understanding that may not be secure. In addition, support materials and resources can be developed that address the common errors and enable students to acquire a better understanding of Mathematics.

References

Confrey, J. (1990) A Review of the Research on Student Conceptions in Mathematics, Science and Programming, *Review of Research in Education*, **16**, 1, 3–56.

Dickson, L., Brown, M., and Gibson, O. (1984) *Children Learning Mathematics: a teacher's guide to recent research*, Oxford: The Alden Press.

Nesher, P. (1987) Towards an Instructional Theory: The Role of Student's Misconceptions, *For the Learning of Mathematics*, **7**, 3, 33–40.

OCR (2012) *Mathematics B (Linear): OCR Report to Centres June 2012*, Cambridge: OCR (Oxford Cambridge and RSA Examinations).

Swan, M. (1990) Becoming numerate: developing conceptual structures. In S. Willis (Ed) *Being numerate: what counts?* Victoria: Australian Council for Educational Research.

Examiners' Reports

OCR GCSE (J512 AND J567)

J512 June 2009 http://www.ocr.org.uk/Images/62540-examiner-s-report-june.pdf

J512 June 2010 http://www.ocr.org.uk/Images/63805-examiner-s-report-june.pdf

J512 June 2011 http://www.ocr.org.uk/Images/63537-examiners-reports-june.pdf

J567 June 2012 http://www.ocr.org.uk/Images/59280-examiners-reports-june.pdf

(All available online; accessed 20 November 2013).

CIE IGCSE (0580)

June 2009 http://teachers.cie.org.uk/docs/dynamic/32880.pdf

June 2010 http://teachers.cie.org.uk/docs/dynamic/39775.pdf

June 2011 http://teachers.cie.org.uk/docs/dynamic/47607.pdf

June 2012 http://teachers.cie.org.uk/docs/dynamic/56160.pdf

(All available online; accessed 20 November 2013).

Question Papers

OCR GCSE (J512 AND J567)

J512 June 2009

Paper 1 http://www.ocr.org.uk/Images/66846-question-paper-unit-j512-01-paper-1-foundation-tier.pdf

Paper 2 http://www.ocr.org.uk/Images/65612-question-paper-unit-j512-02-paper-2-foundation-tier.pdf

Paper 3 http://www.ocr.org.uk/Images/60182-question-paper-unit-j512-03-paper-3-higher-tier.pdf

Paper 4 http://www.ocr.org.uk/Images/65613-question-paper-unit-j512-04-paper-4-higher-tier.pdf

J512 June 2010

Paper 1 http://www.ocr.org.uk/Images/61421-question-paper-unit-j512-01-paper-1-foundation-tier.pdf

Paper 2 http://www.ocr.org.uk/Images/65034-question-paper-unit-j512-02-paper-2-foundation-tier.pdf

Paper 3 http://www.ocr.org.uk/Images/59596-question-paper-unit-j512-03-paper-3-higher-tier.pdf

Paper 4 http://www.ocr.org.uk/Images/63231-question-paper-unit-j512-04-paper-4-higher-tier.pdf

J512 June 2011

Paper 1 http://www.ocr.org.uk/Images/65914-question-paper-unit-j512-01-paper-1-foundation-tier.pdf

Paper 2 http://www.ocr.org.uk/Images/58670-question-paper-unit-j512-02-paper-2-foundation-tier.pdf

Paper 3 http://www.ocr.org.uk/Images/62305-question-paper-unit-j512-03-paper-3-higher-tier.pdf

Paper 4 http://www.ocr.org.uk/Images/65915-question-paper-unit-j512-04-paper-4-higher-tier.pdf

J567 June 2012

Paper 1 http://www.ocr.org.uk/Images/134458-question-paper-unit-j567-01-paper-1-foundation-tier.pdf

Paper 2 http://www.ocr.org.uk/Images/134459-question-paper-unit-j567-02-paper-2-foundation-tier.pdf

Paper 3 http://www.ocr.org.uk/Images/134460-question-paper-unit-j567-03-paper-3-higher-tier.pdf

Paper 4 http://www.ocr.org.uk/Images/134461-question-paper-unit-j567-04paper-4-higher-tier.pdf

(All available online; accessed 27 November 2013).

CIE IGCSE (0580)

June 2009

Paper 1 Paper 2 Paper 3 Paper 4

June 2010

Paper 1 Paper 2 Paper 3 Paper 4

June 2011

Paper 1 Paper 2 Paper 3 Paper 4

June 2012

Paper 1 Paper 2 Paper 3 Paper 4

(All available online from Cambridge International Examinations through their secure *CIE TeacherSupport* website).

Appendix 1: Descriptions of the qualifications

OCR Mathematics A GSCE (J512)

OCR Mathematics A GSCE (J512) was offered until the June 2011 session and was comprised of four papers. Students had to take either two papers at the foundation tier (Paper 1 and Paper 2) or two papers at the higher tier (Paper 3 and Paper 4).

Table 1 GCSE Mathematics papers (J512)

Paper	*Description*	*Calculator*	*Duration*	*Weight*
1	Mathematics Paper 1 – Foundation	No	2 hours	50%
2	Mathematics Paper 2 – Foundation	Yes	2 hours	50%
3	Mathematics Paper 3 – Higher	No	2 hours	50%
4	Mathematics Paper 4 – Higher	Yes	2 hours	50%

OCR Mathematics B GSCE (J567)

OCR Mathematics B GSCE (J567) is comprised of four papers. Students must take either two papers at the foundation tier (Paper 1 and Paper 2) or two papers at the higher tier (Paper 3 and Paper 4).

Table 2 GCSE Mathematics papers (J567)

Paper	*Description*	*Calculator*	*Duration*	*Weight*
1	Mathematics Paper 1 – Foundation	No	1 hour 30 mins	50%
2	Mathematics Paper 2 – Foundation	Yes	1 hour 30 mins	50%
3	Mathematics Paper 3 – Higher	No	1 hour 45 mins	50%
4	Mathematics Paper 4 – Higher	Yes	1 hour 45 mins	50%

CIE Cambridge IGCSE Mathematics (0580)

CIE Cambridge IGCSE Mathematics (0580) is comprised of four papers. Students must take either two core papers (Paper 12 and Paper 32) or two extension papers (Paper 22 and Paper 42).

Table 3 IGCSE Mathematics papers (0580)

Paper	*Description*	*Calculator*	*Duration*	*Weight*
12	Short answer questions – Core curriculum	Yes	1 hour	35%
22	Short answer questions – Extended curriculum	Yes	1 hour 30 mins	35%
32	Structured questions – Core curriculum	Yes	2 hours	65%
42	Structured questions – Extended curriculum	Yes	2 hours 30 mins	65%

Context in Mathematics examination questions

Jackie Greatorex Research Division

Introduction

For at least two decades educationalists have debated whether Mathematics examination questions should be set in context. The advantages of context in Mathematics questions include promoting mathematical literacy, using mathematical tools and thinking to make sense of the world (Debba, 2011; du Feu, 2001). The disadvantages incorporate the difficulties in finding real-life contexts in which school Mathematics readily fits, and therefore examination questions can contain artificial contexts that require learners to make unrealistic assumptions (Clausen-May, 2006; Little and Jones, 2010). This is a contemporary debate. For instance, on 8 October 2013 the Government announced they were supporting schools and colleges to teach *Core Mathematics qualifications* (Department for Education, 2013). Solving significant problems in contexts is at the centre of the Core Mathematics qualifications. Guidance suggests that awarding organisations should assess Core Mathematics in context, using contexts suggested by higher education and industry (Browne *et al*., 2013).

The aim of this article is to revisit the debate to answer the following questions:

1. What are the advantages and disadvantages of examining Mathematics in context?
2. What are the features of a high quality context?

Initially several taxonomies (categories or classification systems) of context are reviewed and the research methods for evaluating the effects of context are considered. Subsequently, the advantages and disadvantages of using context in Mathematics examination questions are explored, focusing on research about public examinations in secondary school Mathematics in England. The literature is used to make recommendations about context in Mathematics questions.

Taxonomies of context

Several authors argue that the term *context* is particularly difficult to define for Mathematics (Berry *et al*., 1999; Little and Jones, 2007; Vappula and Clausen-May, 2006). Berry *et al*. (1999) believe that no definition can be found and consider that it is more useful to think about routine and non-routine questions. Routine questions are those to which learners are likely to respond with a prepared routine consisting of a small number of stages. Non-routine questions do not fit this description.

Whilst there is a dearth of definitions of context, taxonomies describing context abound. Some taxonomies are concerned with Mathematics for learners up to and including 16 years of age (Ahmed and Pollitt, 2007; McCusker, Nicholson, and Ridgway, 2010; Mevarech and Stern, 1997; Vappula and Clausen-May, 2006; Watanabe and Ischinger, 2009) whilst other taxonomies relate to learners aged 17 or 18 years (Debba, 2011; Little and Jones, 2007). The taxonomies are summarised below to further describe context and introduce terms.

Mevarech and Stern (1997) refer to *sparse* versus *real* contexts of questions about linear graphs. Although no clear definitions are provided, their work implies that real contexts are everyday contexts, and sparse contexts are more school-orientated contexts and potentially abstracted from everyday life.

Vappula and Clausen-May (2006) argue that contextualised questions are those that include stories about real-life events, a pictorial or a verbal model.

Watanabe and Ischinger (2009) categorise contexts according to content, including:

- *Personal contexts* – which are of direct personal relevance to learners
- *Educational and occupational contexts* – which include scenarios that learners might contend with while at school, including somewhat artificial problems, or problems that would be encountered in a workplace
- *Public contexts* – which are scenarios experienced in everyday life such as reading part of a newspaper
- *Scientific contexts* – which occur when a question is in a science context such as presenting experimental data.

Debba (2011) uses work by du Feu (2001) to derive question types, according to the characteristics of their context. The categories are:

- *Context-free* – tasks with no context, often simple equations and one-step arithmetic
- *Real contexts* – real problems which mention any named individual(s), institution(s), artefact(s), organism(s) or product(s). These contexts are chiefly statistical. Data are used and quoted and a source is acknowledged
- *Cleaned contexts* – real-life contexts are simplified such that the question is accessible to the learner and suitable for the time constraints of an examination
- *Parables* – fictitious contexts ascribed to an anonymous person/company/organism
- *Contrived contexts* – devised to fit a particular mathematical point, regardless of whether they are relevant to real life.

Some authors also classify questions according to the purpose of the context. Vappula and Clausen-May (2006) argue there are two purposes. The first is getting the story across (explaining the context), rather than supporting the learner with the Mathematics. The second is providing a model for the learner to think with. In this second case, the question and the context start to guide the learner towards a solution.

Watanabe and Ischinger (2009) also offer a classification system which relates to the purpose of the context:

- *Zero order* – the context is not needed to solve the problem
- *First order* – the context is relevant and needed to solve the problem and judge the correctness of the learner's answer
- *Second order* – the learner must engage with both the mathematical problem and its context to solve the problem. Furthermore, to judge the correctness of the learner's answer, the examiner must consider it within the context.

There are similarities between the taxonomies structured according to the purpose of the context. Zero order contexts (Watanabe and Ischinger, 2009) are similar to contexts intended to get the story across and not to support the Mathematics (Vappula and Clausen-May, 2006). First and second order contexts (Watanabe and Ischinger, 2009) are similar to contexts that provide a model for the learner to think with (Vappula and Clausen-May, 2006). A limitation of the taxonomies is that they generally omit marking with the exception of the taxonomy by Watanabe and Ischinger (2009).

The taxonomies summarised so far are descriptive or are structured according to the purpose of contexts. There are four further taxonomies which can be used to evaluate the quality of a context.

Ahmed and Pollitt (2007) classify questions according to whether they are *focused* or *unfocused*. A focused question addresses the aspects of the context that are most salient in real life for the learners. Unfocused questions do not address the aspects of the context that are most salient in real life for learners. A more focused context helps to activate relevant concepts, rather than interfering with comprehension and reasoning.

Little and Jones (2007), Little (2008) and Little (2010) refer to various characteristics of contexts:

- *Accessibility* – the familiarity and comprehensibility of the context, including the comprehensibility of the language and the clarity of the match between a context and a mathematical model
- *Realistic* – the fit of the mathematical model to the real world:
  a. *Natural contexts* – match reality
  b. *Synthetic contexts* – configure reality to match the Mathematics and can reify abstract mathematical ideas
- *Authentic* – the relevance and usefulness of the solution to the real-world context.

Wiliam (1997) reports a taxonomy of context, although his refers to context in the presentation of mathematical texts in general. He lists three types of context in Mathematics teaching:

- *Maths looking for somewhere to happen* – contexts that have little or nothing to do with the Mathematics being taught. The context is to justify the subject matter. This is very similar to synthetic contexts mentioned above
- *Realistic Mathematics* – contexts with a structure that maps to the mathematical structures being taught
- *Real problems* – contexts in which the key aim is solving a problem, and Mathematics may or may not be needed to find a solution.

Finally, McCusker *et al*. (2010) claim that most statistics questions from General Certificate of Secondary Education (GCSE) Statistics/Mathematics can normally be categorised within one of the following categories:

- *Real world but irrelevant or uninteresting* – contexts about the real world which are irrelevant or uninteresting to the majority of learners taking GCSE
- *Real world but not age appropriate* – contexts about the real world which are not age appropriate for the majority of learners taking GCSE. Boaler (1993b) and Debba (2011) also make the point that many contexts are about the adult world as experienced by examination setters, rather than the world as experienced by children
- *Context irrelevant* – contexts which are irrelevant to the real world
- *Unrealistic context* – contexts which present unrealistic information.

This implies that contexts in GCSE questions about statistics are generally flawed. Whilst there are several taxonomies of context, they are broadly similar. For the purpose of this review, context is taken to include all of the taxonomies discussed above.

Methods used to research the effects of context

Researchers use a variety of methods (often in combination) to investigate how context influences learners' behaviours, cognition and marks. The most prevalent design is a mixed methods study combining a quantitative experiment and a qualitative investigation. There are several experimental designs including matched pairs where learners in the control and experimental groups are matched for ability, attempt different versions of questions and all marks are quantitatively analysed. The methods for the qualitative investigation are wide ranging:

- An analysis of learners' responses to questions (Boaler, 1993a; Fisher-Hoch, *et al*. 1997; Vappula and Clausen-May, 2006)
- Interviewing learners about their experience of answering the questions, once they complete the questions (Ahmed and Pollitt, 2001; Crisp *et al*., 2008; Hong-Kim and Goetz, 1994)
- Stimulated recall sessions when learners are played a recording of themselves answering a question and asked to explain how they attempted the question (Ahmed and Pollitt, 2000)
- A questionnaire completed by learners about their experience of answering the question (Clausen-May, 2006; Khateeb, 2008; Little, 2010; Little and Jones, 2010; Song, 2011)
- An analysis of learners' written explanations of how they attempted questions (Cooper and Harries, 2002).

Mixed methods studies also combine quantitative analysis of marks with qualitative analysis of the questions (Watanabe and Ischinger, 2009). Other researchers add to this design by asking learners to complete questionnaires about their experiences (Cresswell and Vayssettes, 2006), or qualitatively analysing answers and interviewing learners (Debba, 2011).

There are fewer studies with a predominantly qualitative approach. Little (2008) and Morgan, Tang, and Sfard (2011) qualitatively analysed examination questions. Cooper (1998) qualitatively analysed responses to examination questions and interviewed learners.

Verbal protocols are a prevalent method of investigating the mental processes used to undertake a task (Kasper, 1998), including testing English as a second language (Green, 1998; Taylor, L., 2005). Verbal protocols are conspicuous by their absence from the list of qualitative methods above. In verbal protocol research, learners provide concurrent 'think aloud' verbal protocols as they answer an examination question (saying what they think as they tackle the problem) and retrospective verbal protocols (explaining their thought processes after completing a task) (Page and Rahimi, 1995; Taylor, K. L. and Dionne, 2000). The verbal protocol data is qualitatively analysed to describe the learners' mental processes. Sometimes the combination of concurrent and retrospective verbal protocols is referred to as *cognitive laboratory* or *cognitive lab* (Levine *et al*., 2002). Cognitive labs are prevalent amongst major testing organisations in the USA (King and Laitusis, 2008; Test Measurement and Research Service, 2010a; Test Measurement and Research Services, 2010b; Zucker, Sassman, and Case, 2004). They are used to investigate a variety of domains: reading, Mathematics, language, spelling, listening, Science and Social Science (Test Measurement and Research Service, 2010a; Test Measurement and Research Services, 2010b; Zucker *et al*., 2004). This suggests that future research about context in GCSE, A levels or equivalent qualifications might benefit from using verbal protocols/cognitive lab.

Examining in context; advantages, disadvantages and recommendations for practice

There are several advantages and disadvantages to testing Mathematics in context, which are reviewed below. According to Brown (1999), examination questions should concentrate on the essence of the school subject, and context is best used in projects when learners have a teacher to guide them. Over a decade later this is still the prevailing expert wisdom. Therefore Browne *et al*. (2013) suggest 50 per cent of the assessment credit in the new core Mathematics qualifications is internal assessment. However, sometimes context is required in examinations. Therefore, suggestions for developing high quality contexts derived from research findings are presented below. They are presented with the caveat that the construct being tested influences what makes a high quality context. For instance, familiar contexts are more accessible to learners and are therefore arguably of high quality. However, if the construct to be tested is whether learners can apply mathematical principles in new situations, then using familiar contexts is invalid. Consequently the suggestions for developing high quality contexts should only be followed if they do not compromise validity.

The following section is organised by assessment topics, such as the construct.

Construct

GCSE Mathematics or A level Mathematics are designed to test school Mathematics as described in specifications, the National Curriculum, Ofqual's subject criteria and other guidance. The extent to which school Mathematics includes any topic or skill may vary over time. In other words, what counts as mathematical knowledge, skills and understanding varies. For the purposes of validity it is important that context does not detract from testing school Mathematics, for instance by encouraging responses using non-mathematical knowledge or creating construct-irrelevant difficulty (non-mathematical difficulty).

Unfortunately, these hazards are found in questions used in some research. Cooper (1998) found that some context questions elicit everyday knowledge or real-world experience rather than mathematical knowledge. Ahmed and Pollitt (2007) found that generally focused questions avoid construct-irrelevant difficulty. Furthermore, context can be successfully used to test whether learners can solve problems in new situations (Ahmed and Pollitt, 2007).

There are aspects of school Mathematics which cannot be tested in examinations with a time limit. For instance, contexts which are sufficiently open for learners to negotiate their own context and develop deep meaningful understandings are too lengthy for an examination, although they can be used for projects (Boaler, 1993b). Examinations cannot test aspects of the modelling cycle, such as discussing assumptions, and refining and critical reading of longer arguments (Little, 2010). Therefore, examination questions (with and without context) are restricted to requiring learners to undertake pseudo modelling. Real-world contexts in examination questions can be perceived as embryonic modelling problems (Little and Jones, 2010).

There are some areas of school Mathematics for which testing in context is clearly suitable. Context in data handling questions provides meaning for the questions; context-free data handling questions may reduce performances as the questions are devoid of meaning (Fisher-Hoch *et al*., 1997).

Therefore, context should:

- generally be used for all data handling questions
- be focused (address the aspects of the context that are most salient in real life for the learners).

Context

The weight of evidence suggests that context in itself is not the main factor making performance generally better or worse. That is context per se is not the main factor influencing the difficulty of questions. Difficulty is quantified in facility values (mean performance on a question, expressed as a percentage of the marks available for the question). Context does not hinder any more learners than do abstract presentations (Vappula and Clausen-May, 2006). Context in questions does not generally advantage learners who experience either of the following two approaches to teaching and learning (Boaler, 1998). One approach is conducting projects with considerable independence until just before the GCSE examination, when learners practise examination questions. The second approach is working through a textbook, asking for help with each new exercise and preparing for examinations. Context does not uniformly alter performance because each learner's individual experience influences how they interpret the context forming their own individual understanding of the context which impacts how they answer the question and how many marks each learner achieves (Boaler, 1993b; Debba, 2011).

If context *per se* is not generally influencing performance and difficulty, then what is? Learners' choices of Mathematics procedures are likely to be determined by the testing situation, rather than by the context (Boaler, 1993b). Competencies, content, item format (complex multiple choice, multiple choice, short answer, multiple short answer and extended response) and word count of an item are all better predictors of difficulty than context (Watanabe and Ischinger, 2009).

Nonetheless, context can shape learners' answers. Contextualisation in division questions encourages pupils to use informal or drawn methods (Vappula and Clausen-May, 2006). When encountering a context-free question part within a contextualised question, many learners fail to move their thinking out of the context and do not realise that the question part requires a simple textbook answer (Ahmed and Pollitt, 2000).

Therefore, context should be applied to all parts of the question.

So far this article has considered the influence of context *per se*. It is possible that context alone does not modify performance, but that aspects of context might shape the responses of individual learners or particular groups of learners. In the remainder of the article the influence that particular aspects of context have on learners' responses and performance are considered.

Realism

There are several advantages to incorporating real-world context in Mathematics examination questions. They reinforce the perception that Mathematics is useful (Little and Jones, 2010). Furthermore, real-world context can help ease the difficulty of questions by providing mental scaffolding for thinking within the context (Little, 2010) and helps learners use context-specific heuristic strategies (Little and Jones, 2010).

However, it is difficult to find real-life contexts in which school Mathematics readily fits (du Feu, 2001). Therefore, many Mathematics examination questions have synthetic contexts, which have disadvantages. Some learners perceive real-world contexts as artificial (Little and Jones, 2010). Context questions often require learners to make unrealistic assumptions (Clausen-May, 2006).

Despite these issues, synthetic contexts can be desirable as they are artificial models of a context which learners can be asked to evaluate. Learners can be asked to discuss and compare models and hence appreciate the relationship between reality and mathematical models. In such questions, synthetic contexts present limited mathematical models (Little, 2010).

Therefore, context should:

- be realistic or ask learners to evaluate synthetic contexts and the relationship between reality and mathematical models.

Relevance/familiarity

The weight of evidence suggests that if the context is unfamiliar or irrelevant to learners' lives then the demand of comprehending the question and the question difficulty increases. Familiar contexts have statistically significant positive effects on learners' inference making, and unfamiliar contexts can create disadvantages for deep comprehension (Song, 2011). Learners from rural poor areas tend to gain few marks on question papers with contexts which are irrelevant to their lives (Vurayai, 2012).

However, there is contradictory evidence on this issue, even from within the same research. For instance, Debba (2011) reports conflicting findings: unfamiliar contexts can be a barrier to performing well, but the level of real-life relevance that learners attribute to a context does not relate to performance.

Context triggers learners' schemas of everyday experience and reasoning which they use to answer questions, and can result in wrong answers (Ahmed and Pollitt, 2000; Debba, 2011). For instance, one learner's free ATM withdrawals influenced their answer to a question asking learners to work out an ATM cash withdrawal fee based on the amount of money withdrawn (Debba, 2011).

When contexts correspond to real-world situations with which learners are familiar, this can cause them to be unsure about whether to answer a question in terms of the subject or everyday knowledge and whether subject knowledge or everyday knowledge is rewarded with marks (Ahmed and Pollitt, 2000). Focused questions generally provoke fewer misunderstandings than unfocused questions (Ahmed and Pollitt, 2007).

Context can help learners answer a question when everyday reasoning and the correct method of answering the question coincided (Debba, 2011). For example, a football fan answered a question about teams and points in a league system. Their experience of the game and the point system helped identify the crucial information for solving the Mathematics in the question (Debba, 2011).

The challenge of writing questions with contexts that are familiar or relevant to learners is that they do not have a uniform experience, nor do they have the same experience as test developers. Therefore, some learners find a context relevant to their real lives, and others do not (Debba, 2011). Furthermore, many contexts represent adult life, or present adult metaphors, as they are written by adults (Boaler, 1993b; Debba, 2011; McCusker *et al*., 2010).

Cresswell and Vayssettes (2006) claim the following are typically encountered in the lives of 15 year olds:

- Tasks involving quantitative, spatial, probabilistic or other mathematical concepts. For example, media outlets (newspapers, magazines, television and the internet) are filled with information in the form of tables, charts and graphs about subjects such as weather, economics, medicine and sports
- Information on issues such as global warming and the greenhouse effect, population growth, oil slicks in the seas, or the disappearing countryside
- The need to read forms, interpret bus and train timetables, carry out transactions involving money and determine the best buy at the market.

McCusker *et al*. (2010), like Cresswell and Vayssettes (2006), advocate using information from the media. Furthermore, McCusker *et al*. (2010) suggest using contexts that are familiar and relevant to learners' lives, as well as context that are interesting to the learners. They suggest taking inspiration for context topics from a list of the most popular books from the learners' age range or their favourite televisions programmes.

In conclusion, context should:

- be focused
- be restricted to those for which the everyday reasoning and the mathematical reasoning used to answer the question are the same
- be relevant to learners' lives and familiar to learners.

Language

The language in questions can change their difficulty (Fisher-Hoch *et al*., 1997). The language describing contexts is sometimes a barrier to understanding the question requirements which can reduce performance, especially amongst learners with low levels of literacy (Debba, 2011). Furthermore, ambiguity in a question can lead to misinterpretations and learners giving the wrong answer (Cooper, 1998). Real-world contexts increase the word length of the questions, and if they are too long they can place too much emphasis on comprehension rather than Mathematics, thus reducing validity (Little, 2010). The word count of a question is a better predictor of difficulty than the type of context (Watanabe and Ischinger, 2009). Since GCSE was first examined in 1988, the readability of Mathematics examination questions has improved on general readability measures, that is the questions are easier to read (Morgan *et al*., 2011). This is due to the examination boards working to avoid obscuring the Mathematics with language (Morgan *et al*., 2011). This reduces the extent to which the Mathematics examinations measure reading ability, and improves the extent to which they assess the intended construct (School Mathematics), thereby improving validity.

Learners expect questions to be written in positive language, and can accept the positive meaning of a statement as correct. Putting negative words in bold like **not** reduces, but does not overcome the problem (Crisp *et al*., 2008).

Therefore, context should:

- be easy to read or have an appropriate reading age
- use positive language
- use clear, concise, unambiguous language which does not obscure the question.

Resources

Resources are diagrams, graphs, illustrations and so on that are used to build a context. Pictorial analogues, rather than real-life contexts provide learners with powerful models to think with which enhance learners' responses (Vappula and Clausen-May, 2006). However, complex contexts, information-heavy contexts and diagrams containing a lot of irrelevant information can all lead to learners producing errors (Debba, 2011).

The resources in a question can prompt learners to use the wrong procedure or method (Boaler, 1993a). Learners generally expect that all resources are necessary to answer the question, and this influences responses to questions. For example, learners can place too much emphasis on a resource (Crisp *et al*., 2008).

Therefore, context should:

- provide pictorial analogues which are models to think with
- use only relevant resources and avoid unnecessary information.

Question layout

Question layout and the order in which information is provided can influence learners' responses. Question layout can shape learners' ability to find crucial information to answer the question (Debba, 2011). Inappropriate ordering of the problem and information within a question increase its difficulty (Fisher-Hoch *et al*., 1997). Question format can also prompt learners to use the wrong procedure or method (Boaler, 1993a).

Therefore, context should:

- present the problem and information in the order in which it is needed to answer the item
- clearly provide the information needed to answer the question.

Marking

The research is dominated by the issue of how context modifies learners' performance and responses. Unusually the taxonomy by Watanabe and Ischinger (2009) illustrates that context impinges on marking as well as answering examination questions. Marking open-ended contextualised questions reputedly requires a good deal of marker training and mark scheme development (Clausen-May, 2006). However, context questions can be designed to measure the intended knowledge, skills and understanding as well as be unchallenging to mark. This is achieved by listing possible solutions to a question, and asking learners to choose the correct solution or provide figures from the solutions (Clausen-May, 2006). Little (2010) reports a view that real-world contexts jeopardise reliability (test-retest and marking consistency). However, Little also found that well designed questions do not have these problems.

Therefore, context should:

- list possible solutions and ask learners to choose the correct solution or provide figures from the solutions.

Conclusions

This review describes taxonomies of context, explores the detail of how context modifies comprehension, performance and difficulty, and provides some suggestions for question writing. This conclusion focuses on wider issues.

Firstly, major testing organisations in the USA are committing a great deal of time and effort to cognitive laboratory and verbal protocol work. There appears to be less research about how learners comprehend and respond to GCSE and A level questions, and certainly less verbal protocol and cognitive laboratory work with learners regarding these examinations. This may reflect different examination development processes. In the USA, many examination questions are pre-tested, whereas in the UK GCSEs and A levels are not generally pre-tested. They are developed based on years of expertise instead. It is recommended that in future research considers verbal protocols and cognitive laboratories as methods of data collection for investigating the A level and GCSE question answering process, or how questions influence the qualitative aspects of learners' responses.

The review highlights that there have been improvements in contextualised questions in recent years. For instance, since the GCSE was introduced, the examination boards worked to avoid obscuring Mathematics with language (Morgan *et al*., 2011). Hence the questions are easier to read according to a number of readability measures (Morgan *et al*., 2011).

How context influences marking is under researched. Research in this area might find ways of improving marking in addition to those proposed by Clausen-May (2006). Furthermore, as new contexts are included in timed written examinations, such as contexts that Higher Education and employers suggest for core Mathematics, new ways of marking may be needed.

It is important to ensure that examination questions with context are of high quality. Several factors are better predictors of performance and difficulty than context. These include competencies, content, item format and word count (Watanabe and Ischinger, 2009). Therefore these factors, along with context, should be attended to in question development. Nonetheless, continued research to identify how context influences comprehension, performance and difficulty is valuable as it facilitates validity and guards against construct-irrelevant difficulty.

References

Ahmed, A., & Pollitt, A. (2000). *Observing context in action*. Paper presented at the International Association for Educational Assessment conference, Jerusalem, Israel.

Ahmed, A., & Pollitt, A. (2001). *Improving the validity of contextualised questions*. Paper presented at the British Educational Research Conference, Leeds. Available online at: http://www.cambridgeassessment.org.uk/Images/109682--improving-the-validity-of-contextualised-questions.pdf (Accessed 28 November 2013).

Ahmed, A., & Pollitt, A. (2007). Improving the quality of contextualized questions: An experimental investigation of focus. *Assessment in Education: Principles, Policy & Practice*, **14**, 2, 201–232.

Berry, J., Maull, W., Johnson, P., & Monaghan, J. (1999). *Routine questions and examination performance*. Paper presented at the Psychology of Mathematics Conference, Haifa, Israel.

Boaler, J. (1993a). Encouraging the transfer of 'school' mathematics to the 'real world' through the integration of process and content, context and culture. *Educational studies in mathematics*, **25**, 4, 341–373.

Boaler, J. (1993b). The Role of Contexts in the Mathematics Classroom: Do they Make Mathematics More" Real"? *For the learning of mathematics*, **13**, 2, 12–17.

Boaler, J. (1998). Alternative approaches to teaching, learning and assessing mathematics. *Evaluation and Program Planning*, **21**, 2, 129–141.

Brown, M. (1999). One mathematics for all? In C. Hoyles, C. Morgan and G. Woodhouse (Eds), *Rethinking the Mathematics Curriculum*, pp.78–90. London: Falmer Press.

Browne, R., Koenig, J., MacKay, N., Sheldon, N., Silcott, N., & Wake, G. (2013). *Report from the expert panel on core mathematics*. Available online at: https://www.acme-uk.org/media/13699/final%2007october2013,%20expert%20panel%20on%20core%20mathematics%20report.pdf (Accessed 28 November 2013).

Clausen-May, T. (2006). Reality and context in maths test questions. *Mathematics in School*, **35**, 5, 9–11.

Cooper, B. (1998). Assessing national curriculum mathematics in England: Exploring children's interpretation of key stage 2 tests in clinical interviews. *Educational studies in mathematics*, **35**, 1, 19–49.

Cooper, B., & Harries, T. (2002). Children's Responses to Contrasting Realistic Mathematics Problems: Just how realistic are children ready to be? *Educational studies in mathematics*, **49**, 1, 1–23.

Cresswell, J., & Vayssettes, S. (2006). Assessing Scientific, Reading and Mathematical Literacy. A framework for PISA.

Crisp, V., Sweiry, E., Ahmed, A., & Pollitt, A. (2008). Tales of the expected: the influence of students' expectations on question validity and implications for writing exam questions. *Educational Research*, **50**, 1, 95–115.

Debba, R. (2011). *An exploration of the strategies used by grade 12 mathematical literacy learners when answering mathematical literacy examination questions based on a variety of real-life contexts*. Master of Education, KwaZulu-Natal. Available online at: http://researchspace.ukzn.ac.za/jspui/bitstream/10413/5814/1/Debba_Rajan_2011.pdf (Accessed 28 November 2012).

Department for Education. (2013). *New maths qualifications to boost numbers studying maths to age 18*. Available online at: https://www.gov.uk/government/news/new-maths-qualifications-to-boost-numbers-studying-maths-to-age-18 (Accessed 11 October 2013).

du Feu, C. (2001). Naming and shaming. *Mathematics in School*, **30**, 3, 2–8.

Fisher-Hoch, H., Hughes, S., & Bramley, T. (1997). *What makes GCSE examination questions difficult? Outcomes of manipulating difficulty of GCSE questions*. Paper presented at the British Educational Research Association, University of York. Available online at: http://www.leeds.ac.uk/educol/documents/000000338.htm (Accessed 28 November 2013).

Green, A. (1998). *Verbal protocol analysis in language testing research: A handbook* (Vol. 5). Cambridge: Cambridge University Press.

Hong-Kim, Y., & Goetz, E. (1994). Context effects on word recognition and reading comprehension of poor and good readers: A test of the interactive-compensatory hypothesis. *Reading Research Quarterly*, **29**, 2, 178–188.

Kasper, G. (1998). Analysing verbal protocols. *Tesol Quarterly*, **32**, 2, 358–362.

Khateeb, M. (2008). *Cognitive Load Theory and Mathematics Education*. University of New South Wales Australia.

King, K. C., & Laitusis, C. C. (2008). *Sample Cognitive Interview Protocol*. Available online at: http://www.ets.org/Media/Research/pdf/Sample_Cognitive_Lab_Interview_Protocol.pdf (Accessed 28 November 2013).

Levine, R., Huberman, M., Buckner, K., & Goldstein, A. A. (2002). *The Measurement of Instructional Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Grade Students and Teachers to Questionnaire Items*. Available online at: http://nces.ed.gov/pubs2002/200206.pdf (Accessed 28 November 2013).

Little, C. (2008). *The functions and effects of real world contextual framing in A/AS pure mathematics questions: developing an evaluative framework*. Paper presented at the British Society for Research into Learning Mathematics, Kings College, London.

Little, C. (2010). *The use of real-world contextual framing in UK university entrance level mathematics examinations*. Doctor of Philosophy, University of Southampton, Southampton. Available online at: http://eprints.soton.ac.uk/166257/ (Accessed 28 November 2013).

Little, C., & Jones, K. (2007). Contexts for pure mathematics: an analysis of A level mathematics papers. *Proceedings of the British Society for Research into Learning Mathematics*, **27**, 1, 48–53.

Little, C., & Jones, K. (2010). *The effect of using real world contexts in post-16 mathematics questions*. Paper presented at the British Congress for Mathematics Education, Manchester University.

McCusker, S., Nicholson, J., & Ridgway, J. (2010). *Statistics assessment: The good, the bad and the ugly*. Paper presented at the Eighth International Conference on Teaching Statistics (ICOTS8), Ljubljana, Slovenia.

Mevarech, Z. R., & Stern, E. (1997). Interaction between knowledge and contexts on understanding abstract mathematical concepts. *Journal of Experimental Child Psychology*, **65**, 1, 68–95.

Morgan, C., Tang, S., & Sfard, A. (2011). *Grammatical structure and mathematical activity: comparing examination questions*. Paper presented at the British Society for Research into Learning Mathematics, Oxford University.

Page, C., & Rahimi, M. (1995). *Concurrent and retrospective verbal protocols in usability testing: Is there value added in collecting both?* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Song, M. (2011). *Effects of Background Context and Signaling on Comprehension Recall and Cognitive Load: The Perspective of Cognitive Load Theory*. Doctor of Philosophy, Nebraska-Lincoln. Available online at: http://digitalcommons.unl.edu/cehsdiss/120/ DigitalCommons@University of Nebraska - Lincoln database. (Accessed 28 November 2013).

Taylor, K. L., & Dionne, J. P. (2000). Accessing problem-solving strategy knowledge: The complementary use of concurrent verbal protocols and retrospective debriefing. *Journal of Educational Psychology*, **92**, 3, 413.

Taylor, L. (2005). Using qualitative research methods in test development and validation. *Research Notes*, **21**, 2–4.

Test Measurement and Research Service. (2010a). *A cognitive lab report for the American Diploma Project Algebra I End of course exam*. Available online at: http://images.pearsonassessments.com/images/tmrs/tmrs_rg/AlgI_Cog_Lab_Report_101410_Final.pdf (Accessed 28 November 2013).

Test Measurement and Research Services. (2010b). *A cognitive lab report for the American Diploma Project Algebra II End of course exam*. Available online at: http://images.pearsonassessments.com/images/tmrs/tmrs_rg/AlgIICogLabReport_101410_Final.pdf (Accessed 28 November 2013).

Vappula, H., & Clausen-May, T. (2006). Context in maths test questions: does it make a difference? *Research in Mathematics Education*, **8**, 1, 99–115.

Vurayai, S. (2012). Poverty penalty in ordinary level mathematics in rural Zimbabwe. *Wudpecker Journal of Educational Research*, **1**, 5, 79–85.

Watanabe, R., & Ischinger, B. (2009). *Learning Mathematics for Life: A Perspective from PISA*. Organisation for Economic Co-operation and Development.

Wiliam, D. (1997). *Relevance as MacGuffin in Mathematics Education*. Paper presented at the British Educational Research Association Conference, University of York. Link to the paper as a Word document available online at: http://www.dylanwiliam.org/Dylan_Wiliams_website/Papers.html (Accessed 28 November 2013).

Zucker, S., Sassman, C., & Case, B. J. (2004). *Cognitive Labs* (Harcourt Technical Report): Harcourt Assessment, Inc.

Method in our madness? The advantages and limitations of mapping other jurisdictions' educational policy and practice

Gill Elliott Research Division

Introduction

Around the world, educational policy makers are looking towards other countries to see what makes them educationally successful. 'High-performing jurisdictions' (HPJ) is a phrase now in fairly common parlance, thanks to the developing strength of the major international comparisons such as PISA, PIRLS and TIMSS and because of the current political interest in such indicators.

The UK, like many other countries, has a deep interest in the educational activities of other nations globally. In September 2012 the UK Secretary of State for Education introduced a reform of the General Certificate of Secondary Education (GCSE).[1] In his statement, which called for the assessment to be replaced, he stressed the importance of comparing England with other nations, particularly those performing strongly upon international comparative measures, and endeavouring to improve the English education system in order to compete with 'the best'; by which he was undoubtedly referring to HPJs. Although the call for the complete replacement of GCSEs has now been withdrawn, their redevelopment is planned and comparisons with policy and practice in other jurisdictions are still strongly advocated politically, not just in respect of assessments, but in all matters educational:

> *I want my children, who are in primary school at the moment, to have the sort of curriculum that children in other countries have, which are doing better than our own.*
>
> Michael Gove, speaking on 'Daybreak', ITV, 8 July 2013.

The art of 'policy borrowing', along with the collation of 'policy wisdom' from other systems requires the collection, collation and interpretation of enormous amounts of information. This is by no means simple. As Marmor *et al*. (2005), writing in the context of US health policies, state:

> *…there is an extraordinary imbalance between the magnitude and speed of the information flows and the capacity to learn useful lessons from them. There is, moreover, a considerable gap between promise and performance in the field of comparative policy studies. Misdirection and superficiality are all too common.*

This article addresses the advantages and limitations of making descriptive comparisons with other jurisdictions. This is often referred to as 'mapping', reflecting the technical definition of the term to mean the construction of graphic representations of information using spatial relationships within the graphic to reveal connections within the data. 'Parallel descriptions' is a term used by Oates (2013) who contrasts this, entirely descriptive, approach with 'analytical transnational comparison' wherein a deeper level of analysis of structure and causes is contained.

These kind of exercises are not unique to any single organisation or country, nor are they a new technique. The Department for Education (DfE) (2012) and the Nuffield Foundation (Hodgen and Pepper, 2010; Hodgen, Marks and Pepper, 2013) both recently used the approach. The DfE used a mapping technique to "analyse the curriculum content of the comparator jurisdictions in order to provide insights into the commonalities and differences", and described the process as "one of the most technically challenging aspects of the content analysis." It seems that the process of carrying out mapping studies and interpreting their findings is sparsely documented. Sumsion and Goodfellow (2004) note that "we found surprisingly little guidance concerning the practicalities of the processes involved."

Given that this type of research is being used to inform thinking in such important contexts, it is vital that the advantages, disadvantages and processes of conducting it are fully reported. This article seeks to take a step in this direction.

What does a mapping exercise/parallel description look like?

Figure 1 shows an example of what a typical mapping exercise may look like. In this instance the columns represent different jurisdictions and the rows pertain to different topics of interest.

These documents are often created using conventional spreadsheet software and can become very large indeed. They are saved from unwieldiness by the facility to hide columns, and the ability to rearrange both rows and columns depending upon how they are being used. In other examples, tables are created using word-processing software, or sets of profiles containing similar information are presented.

The selection of what, exactly, the rows and columns represent is key to the use of the technique as 'mapping', rather than simple recording of information. If, as in the example, different jurisdictions are contained within each column and specific information in the rows, it becomes possible to read across the rows to make direct comparisons between jurisdictions about a particular feature of interest and to read down columns to set the information into each country's context.

Another common technique for systematising information within a mapping study, for example, when mapping curriculum content, is to set the information from a 'master' curriculum into the left hand column (using a new row for each new topic area), and to map all comparators to that master. If this technique is adopted, it must be decided whether topic areas contained in comparator curricula but not in the master

1. GCSEs are taken in a wide range of subjects by the majority of students in England during Year 11 (age 16).

	Jurisdiction A	Jurisdiction B	Jurisdiction C
Cohort size	*2,075,311 students in middle schools in 2006 (middle school covers 3 school years). 99% enrolment in middle schools. 1% of grade 9 was about 6,000 students in 2004.*	*Between 2005 and 2009 the number of leaving certificates (at the end of compulsory education, age 16) ranged between 63,800 and 66,800. Approximately 95,000 registrations are taken for the matriculation examination (at the end of schooling) every year. (This number includes some mature candidates.) In 2011, 12,769 students took A (advanced) level mathematics and 17,799 students took C (basic) level mathematics.*	*Education enrolment by level > Primary level = 7,268,928 Education enrolment by level > Secondary level = 8,131,217 Education enrolment by level > Tertiary level = 3,984,400*
School system structure	*Middle school mostly determined by a lottery (providing a qualification test has been passed). Secondary schooling runs from 12 to 18, with middle schools from 12 to 15 (school years 7-9, sometimes referred to as middle school grades 1-3). Students who have graduated from middle school can attend high school. Students attend the high school in their own district.*	*Compulsory education starts in the year that the child turns 7. There are then 9 years of compulsory schooling. Upper secondary schooling then comprises three years in a choice of either general or vocational programme. An optional year 10 is available to students at the end of compulsory education, who do not feel that they are yet prepared for upper secondary. In 2009, there were a total of nearly 3 100 comprehensive schools, and the network covers the whole country. 45 per cent of the schools accommodate less than 100 pupils. The largest schools have over 900 pupils. During the past few years, the tendency has been to cut down the number of schools through closures, especially of the smallest ones. Since 2004, the total number is down by appr. 500 schools.*	*Has a nine year compulsory comprehensive school system, which runs from 6-15, and a post-compulsory phase from 15-18. Students have to take entrance examination for junior high school, high school and university, if they change institution. It is always the case in public schools and universities.*
Presentation of curriculum (e.g. presented as learning focus and standards, learning outcomes, level statements, content statements)	*There are 5 sections of the 2007 curriculum: Characters: sets out necessity of studying maths and outlines stands of learning. Objectives: sets out objectives of school mathematics, and mathematics at that school level (e.g. middle school). Contents: Learning content of strands set out in 5 areas - numbers and operations, figures, measure, probability and statistics, patterns and problem solving (at middle school it is assigned to particular grades). Teaching & learning methods: considerations teachers concentrate on when teaching principles and rules etc. Evaluation: Ways in which mathematics can be evaluation to help students' learning.*	*Short statement about the core task of instruction Objectives Core contents, arranged into the following themes: -Thinking skills and methods -Numbers and calculations -Algebra -Functions -Geometry -Probability and Statistics Final assessment criteria for a grade of 8 (arranged under the same area as the core content).*	*Objectives and contents for each grade. The terms and symbols that are learned in that grade are also included. There are also comments about writing lesson plans, and the content that is taught in each grade.*

Figure 1: Extract from a mapping exercise

curriculum are to be recorded or not. If they are, extra rows will be used at the bottom of the table, which remain blank for the master curriculum.

Where the spreadsheets are not too large, it is possible to carry out a supplementary exercise, summarising the findings at the end of each row and column. This allows a picture to emerge of the range of policy and practice in existence, relating to each of the issues investigated (from the rows) and an overall picture of each jurisdiction (from the columns).

The finished mapping exercises themselves are a tool, rather than a 'result' *per se*, which can be used by suitable experts to inform their thinking. The spreadsheets can become very large, but remain easily useable, as users tend to home in on a particular topic or jurisdiction and then broaden their use of the spreadsheet horizontally (other jurisdictions' approaches to the same issue or content) or vertically (contextualisation of the issue or content within the jurisdiction of origin).

In some cases, this method of systematising the information is not sufficient. Cambridge Assessment recently experimented with using separate spreadsheets for three aspects of the education system – the system level (how schooling is organised, age of entry, etc.), the curriculum level (what content is taught at which ages) and the assessment level (what is assessed, when and for what purpose). This worked well; the separate exercises could be cross-referenced when necessary, and the 'layered' approach facilitated cross-level analysis of the information.

The key to successful construction of mapping documents lies in the careful selection of the exact format to be adopted, the material which will be covered, and the jurisdictions or countries to be featured. Resourcing is a not inconsiderable concern. Accurate mapping documents are time-consuming to prepare and often require some subject specialised knowledge. A simple study comparing, for example, six countries on about twenty features of their examination system (or mapping the curricula of a single subject) is likely to cost in the region of £3,000 at current rates simply due to the number of hours work required from a suitably qualified person. If many more features of the system are to be compared (and it will be argued later that a comprehensive study needs all possible aspects of the system to be investigated) or multiple subjects are required, costs can easily increase tenfold. If original documents are not available in the language of the researcher, translation costs will add substantially to this.

Sourcing information

Sourcing information for mapping documents can be tricky. In the case of, say, a comparison of curricula, the information may be relatively self-contained and, as long as it is possible to obtain the correct documents (see the discussion of limitations later in this paper); no further difficulties may be encountered. However, if one is embarking upon a mapping exercise where different pieces of information can be sourced in different ways, there is a clear hierarchy of available sources, each of which has advantages and disadvantages.

Provenance and veracity of information is clearly crucial, whilst proximity to source is also important – primary sources are often considered better then secondary sources when seeking facts, although this has been contested (Barton, 2005), and there are a number of highly reputed international comparative organisations whose information is likely to be as accurate as any primary source (and more accurate than some). Figure 2 shows a notional relationship of each of six key sources of information to both provenance and proximity. Whilst there may well be individual exceptions to these rules, in general these reflect our experiences.

Practical and methodological dilemmas

The issue of how much information to record can become problematic for a conscientious researcher. In some instances, national documents or examples of curricula may contain paragraphs of detail, others just a few words. If the largest amounts of available information are recorded in full, then the document quickly becomes very large; if there are numerous instances of very succinct information combined with a few large pieces, the blank spaces around the shorter information can be distracting at best and infuriating at worst to work with. Ever-decreasing font size and frequent revisiting of source material to adjust the level of detail become stock-in-trade tools for those conducting these studies.

Allied to this is the problem of whether to seek to record source material in its own words or whether to summarise or expedite the material in some way. Recording the material in its original form has the advantage of accuracy, but can be very wordy. Additionally, differences in writing style between the original sources of information can result in a very disjointed document which is difficult to read. Summarising the information into the researcher's words can introduce accidental changes in emphasis, meaning and content. Other strategies such as recording segments of the original with ellipsis or introducing tables or bulleted points where prose existed may inadvertently change how the information is interpreted. One solution (and we would advocate this as good practice) is to provide a short summary in the researcher's own words, plus a web link to the original, more detailed source of information. This has the additional advantage of ensuring that anyone using the document has ready access to relevant primary source material.

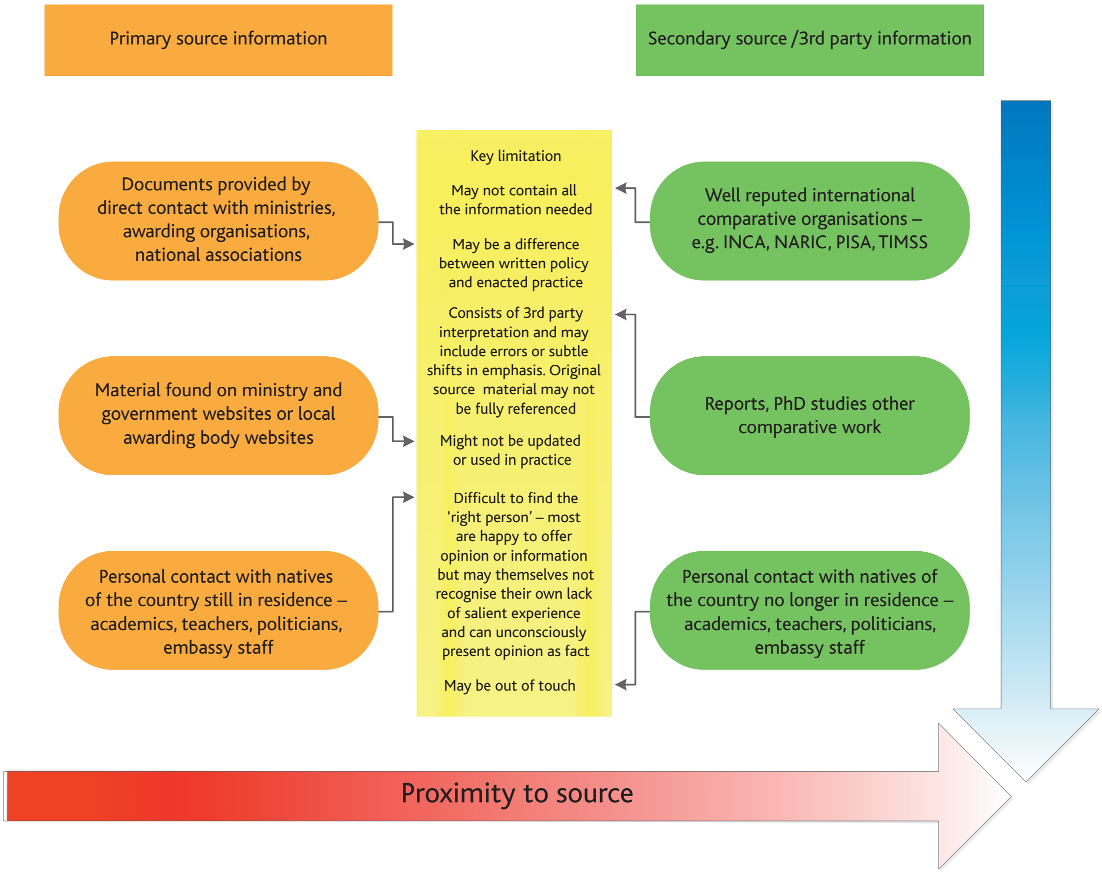


Figure 2: Sourcing information for mapping comparative policies and practice

Of course, it does not work so well when the original material is hard copy, although it might be scanned and added as appendices.

As stated earlier, the finished mapping exercises themselves are a tool which can be used by suitable experts to inform their thinking. Consideration of who is going to use the document, and how they are going to use it, is vital when planning the research. If the document is to be published widely anyone, irrespective of expertise, will be able to draw conclusions from it. In some cases, for example, when users of the document are likely to need specialised expertise in order to interpret the information correctly, this would mitigate against publication, or at least indicate the need to publish in a suitable specialist environment. If the information in the document is likely to be fallen upon by the media or in the political arena, such decisions will need handling with particular care in order that the information within is not only robust but also resistant to misinterpretation.

In many instances the mapping document is for use by the research team themselves or by subject or other suitable experts. In these instances decisions about the level of detail in the content, and also the nature of any limitations encountered when populating the mapping document, will be relatively easy to communicate between different members of the team and the dangers of over-claiming the results or misinterpreting grey areas is greatly reduced.

Strengths and limitations of mapping exercises

It is important to consider the technical strengths and the weaknesses of this approach before embarking upon it, and these are summarised in Figure 3.

As well as considering the advantages and disadvantages of individual studies, it is worth noting just how many such studies, large and small, are undertaken in separate institutions. Few are published formally, and rightly so. In most cases, there are a number of reasons why publication is unwise:

- The content is not absolutely complete
- The accuracy of the content cannot be satisfactorily verified
- The purpose for which the material has been amassed is specific to a particular research question and inappropriate for more general use
- Material would be out of date by the time of any publication

Nevertheless, the sheer volume, and commensurate expense, of this type of work – and there is evidence that it is occurring in similar measure throughout the developed world – should not be underestimated. It is probable that, globally, a huge amount of money is being spent on a research technique that may be lacking in important areas.

A particular danger is that the accessibility of some information leads to a conviction that such comparisons are strong. For example, in February 2013, the UK Education Select Committee travelled to Singapore for four days and one of its members, Craig Whittaker, MP, announced on his web log that: "We quickly started to understand how Singapore produce the best results in the world…"[2]

It is difficult to believe that a complete understanding of such a complex area could occur in just four days. To be fair, the text only says 'started to understand', but there is a clear sense of excitement and persuasion about what had been seen and heard. However, without complete understanding of all pertinent information, these comparisons are necessarily extremely limited.

2. http://craigwhittakermp.wordpress.com/2012/04/23/education-select-committee-visit-to-singapore/, (Accessed 12 July 2013).

Figure 3: A summary of strengths and limitations of mapping exercises as a method of investigating policy and practice elsewhere

Strengths of mapping exercises as a method of investigating policy and practice elsewhere

- Mapping per se facilitates an overview of different jurisdictions with relative efficiency.
- Different 'layers' of mapping documents can provide an effective way of examining the whole system. For example, system level, curriculum level and assessment level. Equally when different subjects are investigated at the same level for a number of jurisdictions, if each mapping document follows the same format, then cross-subject commonalities and differences may also be readily identified.
- Mapping exercises can be extended ad infinitum when required. Extra jurisdictions, or extra areas for investigation, can be added.
- Parallel descriptions can (and should) be updated regularly, if the document is to exist as an on-going resource, otherwise it will become outdated. However, previous versions can be kept as a snapshot of the time in question. For example, the INCA (International Review of Curriculum and Assessment Frameworks) was actively maintained by NFER (supported by QCDA and then QCA) between 1996 and 2013. Upon cessation of active maintenance, a snapshot was taken for posterity by the UK National Archives.

Limitations of mapping exercises

- There is likely to be information which you know must exist, but cannot be obtained.
- There is likely to be information which you have sourced but may in fact be misleading. For example, the written versus enacted curriculum may differ.
- There will be information which you don't know you should even be looking for. Something, perhaps so different to your own culture that you would not think of it.
- Written information (e.g. curriculum material) is often substantiated and exemplified by additional documentation which is not included within the principal record. For example, a curriculum document may have associated schemes of work, and the detail of some jurisdictions' curricula may be contained within state-regulated textbooks. Failure to source or to appreciate the importance of such additional materials will, inevitably, produce extremely misleading results.
- One-off mapping exercises provide a snapshot in time – that time being when the source information was valid (not necessarily when the mapping exercise was completed).
- Policies change – the success of a particular group of students on an international comparative test such as PISA may be due to previous, now outdated, policy. Identifying the appropriate materials to map in this situation can be difficult, and obtaining non-current documentation even harder. This could be termed the 'time-shift problem'.
- Whilst you believe an education system to be good, because that jurisdiction is an HPJ, they themselves may be dissatisfied with the system and be looking elsewhere for inspiration.
- If documents need to be translated, there can be some uncertainty about the accuracy of the translation. Nuances of language can change meaning in very technical ways – a professional, educationally focussed, translation service will be required. Even before any such professional service can be used, some identification of the appropriate material to be translated must occur, and it is very difficult to source material in an unfamiliar language. The chances of finding all the right documents are really quite slim, especially when the third and fourth point above are taken into account.
- The apparent sophistication of large mapping exercises sometimes belies the fact that it simply may not be sensible to be making those comparisons in the first place.

Transforming a parallel description into an analytical transnational comparison

Mapping studies are essentially a systematised method of providing parallel descriptions of policy and practice across different jurisdictions. Parallel description is useful in identifying examples of good practice to follow and poor practice to avoid. It can be used to explore the infrastructure surrounding particular features of systems. It can highlight a variety of approaches and also illuminate practices which are common to many jurisdictions. It can also lead to the realisation that a part of your system might be substantially improved, but it does not provide the evidence necessary to justify major changes in policy or practice (policy borrowing), as Marmor *et al*. (2005) argued when discussing health services conferences in 1990s America:

> *Understood as simply wanting to stretch one's mind – to explore what is possible conceptually, or what others have managed to achieve – this is unexceptionable. Understood as the pursuit of the best model, absent further exploration of the political, social, and economic context required for implementation, this is wishful thinking.*
>
> (Marmor *et al*. 2005).

Effective parallel description is arguably essential to analytical transnational comparison. Certainly it is a highly desirable precursor. Accurate and complete information is clearly fundamental to this further stage of the process. However, analytical transnational comparison requires, at least, two further factors:

1. Much wider and additional contextual information derived from a wide range of sources – social, political, historical, cultural, economic, and educational – all of which interact in a unique dynamic in each individual jurisdiction.
2. A team of information analysts with expertise across all the areas listed in (1) to interpret the material effectively.

Discussion

Often, mapping exercises are used to support investigations into HPJs, as identified by international comparative studies like PISA or TIMSS. In this context, mapping exercises, which provide parallel descriptions of the jurisdictions being compared, can help interpret a situation where ranked position is only a part of the picture. In a further stage, if both the parallel descriptions and the ranked positions on international comparative studies are used as sources of evidence and deeper insights are sought into the reasons why particular strategies succeed in certain places, then valuable intelligence can be developed which might warrant the term 'policy wisdom'.

It would be dangerous to embark upon a comparative approach, without a clear vision of the limitations involved. However carefully mapping documents are constructed, there can be issues with the construction, interpretation and use of the information.

All too often, principal written records are substantiated, exemplified or modified by additional documentation or even oral evidence. For example, the published mark scheme will not have been used in practice without standardisation procedures and communications between different members of the marking team, little of which may be evidenced in the public record.

It is not always possible to source all information; publicly available documents may not contain the answers to all the questions, and if you contact individuals you cannot always be sure that they are the people best placed to answer the query (even if they themselves think that they are). The intended, and documented, curriculum may be considerably different from the enacted curriculum. By its very nature, omitted information can dramatically skew the picture you receive and your interpretation of that picture. For example, researchers working from documents on the web, or even visiting the jurisdiction in question, are unlikely to see the full picture of education in that jurisdiction. Less successful schools, or elements of the system, are not likely to be shown off in public, by either high- or low-performing jurisdictions.

There is a time lag in the findings of major international studies such as PISA, TIMSS and PIRLS. Success in such studies is most likely related to policies and practices which occurred some years before the studies themselves were conducted, and even further before the results were made available. Allied to this is the fact that most jurisdictions' policies are in some state of evolution or flux most of the time, and few jurisdictions, however successful, are content with their current performance. Identifying the policies and practices which contribute to the success is like catching fireflies – there are a huge number of tiny factors which influence the big picture, and they are gone before you can step towards them.

Misleading information can emerge if a structured approach to comparison is not followed. For example, if a great deal is known about the assessment structure of one jurisdiction and little about the school system, but the opposite applies to another jurisdiction, then comparisons between them will be, at best, patchy. It is also human nature for individuals to become inspired by a particular approach encountered, and potentially blinkered when viewing alternatives. A systematic, carefully constructed, rigorous foundation for comparison, such as a well-conducted parallel description can mitigate against this.

Cause and effect can be readily confused. For example, one could attribute success to a particular element of the system common to HPJs, simply because it was common to a number of them. However, if that same element is also present within the systems of low-performing jurisdictions (and these are rarely investigated to check) it cannot simply be the existence of the policy or practice which is the sole cause for success, and attribution for success must be sought in the details of how it is used.

Educational policies and practice do not exist in isolation. There is a whole web of inter-related circumstances which contribute to the success (or otherwise) of any educational policy – overall culture, parental expectations, dynamics within and outside schools, teen and youth culture, attitudes to teaching and learning, economic performance of the country with its concomitant effect upon disposable income, family attitudes and motivation. This is where analytical transnational comparison comes in.

To plaudit elements of alternative systems without having a clear view of how those elements sit within the context of that system is unlikely to prove fruitful. For example, it might be the case that a very successful jurisdiction sets challenging compulsory examinations at age 15 and students perform well on them, but do so within the context of streaming candidates from an extremely young age and investing very heavily in support for the lower-performing students. To adopt the immediate finding which pertains to the age group of students we are most interested in (challenging examinations at age 15) without pairing it with the information about the approach followed at an earlier stage would be mistaken. It also courts the danger of imposing unsuitable elements into the UK system which are unlikely to be successful in the long term.

Nevertheless, there is much that can be gleaned from studying other jurisdictions' approaches to education if comparisons are undertaken in a pragmatic and systematic way. It is crucial to trace the full picture about alternative or innovative approaches – where they are used, how they are used, and upon what other elements of the system they are interdependent. Are such approaches directly linked to success, or are there confounding factors? Cross-referencing between different jurisdictions can be illuminating, especially if lower-performing jurisdictions are also considered. If the same policy is followed, is it accompanied by similar practice in other aspects of the educational system? How does it work here, but not there?

With sound methodological practice this type of study can contribute to the debate about educational reform, but without it the results can be extremely damaging.

References

Department for Education (2012). *Review of the National Curriculum in England. What can we learn from the English, mathematics and science curricula of high- performing jurisdictions?* Research Report DFE-RR178. Available online at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/184064/DFE-RR178.pdf (Accessed 26 November 2013).

Barton, K. C. (2005) Primary Sources in History: Breaking through the Myths. *The Phi Delta Kappan*, **86**, 10, 745–753. Available online at: http://ldc.wiki.hempfieldsd.org/file/view/Article_Primary_Sources_in_History__Breaking_Through_the_Myths.pdf/424625974/Article_Primary_Sources_in_History__Breaking_Through_the_Myths.pdf (Accessed 26 November 2013).

Hodgen, J., Marks, R. and Pepper, D. (2013) *Towards universal participation in post-16 mathematics: lessons from high-performing countries*. Nuffield Foundation. Available online at:https://www.primarydandt.org/res/documents/page/Towards_universal_participation_in_post_16_maths_v_FINAL.pdf (Accessed 26 July 2013).

Hodgen, J.and Pepper, D. (2010) Is the UK an outlier? An international comparison of upper secondary mathematics education. Nuffield Foundation. Available online at:http://www.nuffieldfoundation.org/sites/default/files/files/Is%20the%20UK%20an%20Outlier_Nuffield%20Foundation_v_FINAL.pdf (Accessed 26 July 2013).

Marmor, T., Freeman, R. & Okma, K. (2005) Comparative perspectives and policy learning in the world of health care. *Journal of comparative policy analysis: Research and Practice*. **7**, 4, 331–348.

Oates, T. (2013) *Why Grand Theory and detailed narrative are equally essential in drawing from transnational comparisons*. Paper presented at the AEA-Europe Conference, 7–9 November 2013, Paris.

Sumsion, J. & Goodfellow, J. (2004): Identifying generic skills through curriculum mapping: a critical evaluation. *Higher Education Research & Development*. **23**, 3, 329–346. Available online at: http://www.tandfonline.com/doi/pdf/10.1080/0729436042000235436 (Accessed 26 November 2013).

Cultural and societal factors in high-performing jurisdictions

Victoria Crisp Research Division

Introduction

This article aims to provide insights into some of the cultural and societal contextual factors that influence education systems, using a number of high-performing jurisdictions (HPJs) as case studies. Consideration of the education and assessment systems of HPJs around the world has become a strategy of some interest during education reform and/or development. However, it has been noted that when doing so, societal and cultural features of the jurisdictions need to be considered (e.g. Elliott and Phuong-Mai, 2008; Alexander, 2010; Oates, 2010; Barber, Donnelly and Rizvi, 2012). The effects of a particular educational system may well be influenced by such factors, and as a result the system of one jurisdiction will not necessarily transfer the educational and achievement benefits if simply replicated in the jurisdiction undergoing change.

This article has been written using various secondary sources such as relevant articles, books and reports, newspaper articles, blog posts and other online material. A number of researchers have previously summarised and analysed the features of HPJs, including some of the cultural factors, to identify the possible reasons for the high achievements of students (at least on some of the measures that have been influential, such as PISA, TIMSS and PIRLS). Such work, key examples being the work of the Center on International Education Benchmarking and the Organisation for Economic Development (OECD) produced book *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education*, was particularly useful to the current article.

Six jurisdictions were chosen as the focus for this exploration of cultural and societal factors. The focus jurisdictions were: Alberta (Canada), Shanghai (China), Hong Kong, Singapore, Victoria (Australia), and New Zealand. A few additional jurisdictions for which cultural issues of interest were also noted during the literature review for this article are also mentioned briefly.

A section for each jurisdiction will now be presented, summarising known aspects of the culture and society that relate to education. These include themes identified by previous explorations of HPJs. After these 'case study' sections, a final section brings together some of the cultural and societal themes that appear to be common to some of the jurisdictions of interest.

Alberta, Canada

Canadian students perform well in international comparisons regardless of their socio-economic status, first language or if they are a recent immigrant (Mehta and Schwartz, 2011). Alberta is one of the higher performing and richer Canadian states, and there are fewer 'new Canadians' in Alberta than in some other states. For example, Ontario has more immigrants and special efforts are made to support language learners for this reason *(ibid)*.

The success of Alberta in international tests is attributed by some to the competition within their education sector (Flanagan, 2011). There is competition between public (secular) schools and public Catholic schools and there is competition between public schools and private schools (religious and secular) *(ibid)*. Private schools are 60 per cent or 70 per cent funded by the Government and supplemented by fees, which makes private education fairly affordable *(ibid)*. (There are also some private schools not subsidised by the Government that follow special programmes rather than using the Alberta curriculum *(ibid)*.) Additionally, there are charter schools founded by groups of parents; these are public schools but may charge fees for extras such as additional tuition *(ibid)*. With these various options and private school being a relatively affordable option, public schools have acted to compete to maintain their student numbers by improving what they offer via various strategies such as introducing special programs *(ibid)*.

There is a culture of inclusiveness and cooperation in Canada, and a focus on peace, order and good government as collective concerns (Gaffield, 2012). Thus education is seen in terms of the impact on society, and is allocated substantial public funds *(ibid)*. There is a generally shared view in Canada that society is collectively responsible for educating all its children (Mehta and Schwartz, 2011). Within this context, teachers feel they have an obligation and responsibility to ensure that all students are educated. A news article described Canada's educational culture as follows:

> *Being Canadian is … about being cooperative and inclusive and about valuing shared community and public life. It's not this or that province's policy that makes Canada such a strong educational performer, but a social fabric that values education and teachers, prizes the public good, and doesn't abandon the weak in its efforts to become economically stronger.* (Hargreaves, 2011).

Education in Canada is decentralised, with each provincial or territorial government responsible for education policy and curriculum development (Center on International Benchmarking [CIEB], 2013a). However, the importance of high standards and best practice is recognised and the Ministries of Education in each province tend to look to one another when making policy decisions (partly through collaboration via the council of Ministers of Education) and thus there are similarities in policies *(ibid)*. School boards oversee the running of clusters of schools (the clusters are based on area but also school type, e.g. religious). The system of school boards was inherited from the USA, via immigrants moving to Canada in the 1700s (Peters, 2011). Local matters were dealt with locally, with those in local towns and villages running schools, though State Government might offer support *(ibid)*.

Alberta saw many educational reforms in the 20th century due to changes in government and changes in the educational practices seen as in vogue (e.g. progressive education was introduced and later abandoned. See Ell, 2002, or Matsumoto, 2002, for more historical detail). In the 1990s some teachers and others expressed concerns about the constant reform, and parents complained about the quality of education their children received (Matsumoto, 2002). There were also various challenges around budget cuts for education in the 1990s *(ibid)*.

Most provinces have a policy of high selectivity in teacher education programmes (CIEB, 2013a). In Canada, teachers are valued, they must undertake a professional program of university-based training, and working conditions for teachers are good (e.g. acceptable pay, good availability of professional development) with teachers given discretion to make professional judgements (Hargreaves, 2011). In Alberta, the Government has funded a programme to support innovation projects in schools, and thus it could be described as allowing a 'bottom-up' approach *(ibid)*.

Canada scored well above average for reading in the PIRLS 2011 results and the degree to which parents engage their children in literacy activities was high (Bradshaw, 2012). Around 70 per cent often read and talk about books before their children start school and children with this experience scored better in PIRLS than those without, suggesting a connection *(ibid)*. Canadian children are more likely than others to read daily for pleasure (Tibbetts, 2007, cited in Mehta and Schwartz, 2011). Parents are supportive of their child's education and are seen as an asset to schools (Mehta and Schwartz, 2011).

There is a strong culture of internal assessment in Canada and there is trust in teachers' abilities to make assessment judgements. Alberta is unusual in that exams determine 50 per cent of high school Diploma grades (Alberta Education, 2013a) – only Alberta and Quebec use exams as part of Diploma grades and Alberta places more weight on them than Quebec (Anonymous, 2011). Some feel 50 per cent weighting of exams is too much and question the validity of some of the tests, for example, questioning the appropriateness of multiple choice questions for some subjects where they are used *(ibid)*. The relatively high use of exams in Alberta is intended to prevent bias and inflation of marks, and provide standardisation *(ibid)*. Interestingly given the high results in international tests, a public survey across Canada by the Canadian Education Association (CEA) found that respondents in Alberta were most likely to say that the education system needs significant change (64 per cent) (CEA, 2012). Whether this relates to dissatisfaction with the use of external assessment and this being culturally more unusual is difficult to say. The 50 per cent weighting of external exams in final Diploma outcomes has been in place since 1983 (Matsumoto, 2002), so is hardly a new feature of the educational landscape in Alberta.

There is some controversy about the grading of student assignments. School boards, and even individual teachers, can make their own decisions about how missing work is scored (Slobodian, 2012). It is not uncommon for a missed assignment to be indicated as 'not handed in' without this lowering a student's overall grade, thus students could choose to 'play the system' and skip work without penalty *(ibid)*. There are calls for the Government to create uniform policies on this issue *(ibid)*.

Shanghai, China

In China, education is traditionally highly valued (Elliott and Phuong-Mai, 2008; Cheng, 2011). From 603 AD to 1905 the Civil Examination system was used to select officials to government jobs (Cheng, 2011). This system was highly competitive and hence contributed to attitudes around the importance of learning *(ibid)*. The exams involved writing essays of political relevance, with reference to the Chinese classics texts (the *Four Books and Five Classics*); hence there was an emphasis on rote-learning of these texts *(ibid)*.

Over time, stories about poor scholars who endured hardship but achieved success in the Civil Examination became part of folklore (Cheng, 2011; CIEB, 2013b), and it was indeed the case that a large percentage of those passing the national exams were from ordinary families (International Qualifications Assessment Service (IQAS), 2007). Thus, the system was seen as encouraging upward mobility (Cheng, 2011). Within this setting, this has led to parents having high aspirations for their children (or at least for their sons) and to an ingrained view that hard work and putting up with hardship is the route to success (Cheng, 2011). This relates to the common Chinese belief that '*diligence can compensate for stupidity*', or in other words, that it is effort and hard work that determine success not innate ability (Cheng, 2011; Ellis and Bratu, 2011; CIEB, 2013b). Chinese culture has also been influenced by the Civil Examination in that education is very focused on examination results as the only way to validate learning and as the only route to upward social mobility (Cheng, 2011). Therefore 'education' as a term in China is synonymous with 'exam preparation' *(ibid)*. The perceived importance of exams has transferred to the contemporary context, and private tuition and attendance at tutorial schools is common *(ibid)*. Confucian philosophy claims that proper education is important to social harmony and that education should be available to all, not just the privilege of an elite few (IQAS, 2007). The goals of education are seen as about realising an individual's potential and discouraging unethical behaviour *(ibid)*.

In Chinese society, families look to their children for support in old age (CIEB, 2013b). This is a factor in families supporting students in their educational efforts and having high expectations of them *(ibid)*. Within this setting teaching has become a high status occupation and students are willing to put significant effort and time into their studies *(ibid)*.

The People's Republic of China was established in 1949 and in the decade that followed a national curriculum and teaching materials were introduced (Soviet model) (Tan, 2012). From 1966 to 1976 China experienced the Cultural Revolution – an effort by government to prevent inequality and stop those with privilege passing this on (Cheng, 2011; Hays, 2012; Tan, 2012). Universities and conventional schools were closed (Cheng, 2011; Tan, 2012). New schools were run by workers, peasants and soldiers, and academics and those with some degree of education or privilege were sent to villages, rural areas or factories for 're-education' (Cheng, 2011; Tan, 2012). Some educated youngsters who had been moved to rural areas during the Cultural Revolution had become teachers during this time, although they did not have teacher training (Cheng, 2011; CIEB, 2013b). The Cultural Revolution left China's education system and curriculum in need of being totally rebuilt in the late 1970s/early 1980s, and with a lack of trained teachers (Cheng, 2011). The widely shared belief in the importance of education, and the view that effort is more important than inherited intelligence, were assets in the aftermath of the Cultural Revolution (CIEB, 2013b). A programme of retraining for under-qualified teachers was instigated after the revolution, though one of the challenges was that once trained many teachers gravitated towards the towns and cities, leaving village schools with even less expertise *(ibid)*. New teachers were attracted by offering priority admissions to universities *(ibid)*. The Teachers Law was passed in 1993 to prescribe minimum educational qualifications for different levels of teaching (IQAS, 2007). By 2010 new primary teachers must have a *zhuanke* qualification from a post-secondary institution and new junior secondary teachers must have a *benke* qualification from a teachers' college, normal university or other relevant Higher Education Institution (HEI) *(ibid)*. Some senior secondary teachers have graduate education from a university *(ibid)*. Thus the qualifications of teachers have dramatically increased since the Cultural Revolution. In addition to initial teacher training, all teachers are required to engage in 240 hours of professional development activities over a five year period (Cheng, 2011).

Raising teacher pay and upgrading teacher education has been an important element of recent education reform in Shanghai (and also in Hong Kong) (Cheng, 2011). Teacher salaries in towns and cities are seen as providing a stable income and although not high, they have increased and there are ways in which teachers can supplement their incomes (CIEB, 2013b). This has increased the desirability of teaching as an occupation (Cheng, 2011). Almost all officials in government education offices were originally teachers, meaning that teaching is also seen as a route into government *(ibid)*.

Teaching is organised centrally (Cheng, 2011) and lesson plans are drafted by teaching study groups (CIEB, 2013b), each of which is supervised by the teaching study office of the relevant Education Bureau, and in turn by the municipal or provincial government (Cheng, 2011). Many hours go into lesson planning, and teachers are regularly observed giving planned lessons (CIEB, 2013b). More senior teachers demonstrate lessons as part of professional development for other teachers *(ibid)*.

Another element of the system is an initiative ('empowered administration') that pairs a weaker and stronger performing school to help the weaker school (CIEB, 2013b). A management team is sent into the weaker school, and strategies include good teachers demonstrating good lessons based on a lesson plan *(ibid)*.

Shanghai has a certain amount of autonomy from China as a whole with regard to education and from the mid-1980s was permitted to set its own entrance exams for HEIs (CIEB, 2013b). This prompted changes to the exams and curriculum away from memorisation (and multiple choice questions) towards problem-solving and drawing on deeper understanding *(ibid)*. Ongoing curriculum reforms since 1998 have focused on preparing schools to meet the needs of economic developments in China (Tan, 2012). One aim is to change the learning style away from an exam-focused knowledge transmission approach towards:

- increased real-life experience;
- teamwork;
- problem solving;
- exploration;
- critical and creative thinking;
- learning being more independent and student-directed *(ibid)*.

The slogan 'return class time to students' is being used to reduce the time spent with teachers' lecturing and increase time allocated to student activities (Cheng, 2011). Subjects are grouped as 'Foundational Subjects', 'Expanded Subjects' and 'Inquiry/Research Subjects' to promote more diverse skills and experiences (Tan, 2012). However, only the Foundational Subjects are assessed by exam and the dominant exam-oriented culture and pressure means that parents, students and teachers still see the exams, and thus the Foundational Subjects, as most important *(ibid)*. This may be limiting the success of the effort to change the kind of learning taking place *(ibid)*. Some educators comment that examination pressure still prevails as examination scores are seen as the most scientific and fair basis for decisions (Cheng, 2011). Many still consider students in Shanghai to lack independence, creativity and innovation (Hays, 2012; Tan, 2012; CIEB, 2013b) and some argue that children are rarely left to learn in a way of their choice, are 'spoon fed' their learning, and thus they have not learned how to learn (Cheng, 2011).

Whilst China has a top-down education system, with the Ministry setting out instructions to local bureaucrats, which are passed down to school administrators and then to teachers, some effort to devolve education from the central to local level has been part of the educational reform policy post Cultural Revolution (Hays, 2012). Indeed, the reform in Shanghai has been positively affected by a number of 'bottom-up' initiatives, but with control and intervention from the municipal government as considered necessary (Cheng, 2011). Another element of the 'control' comes from the Chinese Communist Party; whilst the Government has authority over the education system, the Communist Party has played a management role since 1949 (Hays, 2012). There is a designated Communist Party secretary in every school, often in a management role, who is responsible for guiding their school in line with party policy *(ibid)*.

Society is perceived as a vertical hierarchy with parents wanting their students to be first in the class and to achieve 100 per cent in results (Cheng, 2011). Parents often have perceived rankings of HEIs and want their students to go to the best one *(ibid)*. Shanghai is a popular location for students across the country to come for Higher Education *(ibid)*. This has increased the sense of competition despite a generous quota for local students *(ibid)*.

As part of China's heritage, students' belief in needing to work hard to succeed leads to a high level of engagement in learning, with students in Shanghai typically fully attentive and engaged during lessons (Ellis and Bratu, 2011). This is reinforced by a lack of tolerance by teachers of non-attention and the cultural expectation that students must concentrate (Phillips, 2010). However, the motivation to study amongst Chinese youngsters is thought to be primarily extrinsic, driven by family and wider expectations rather than by intrinsic interest (Cheng, 2011).

Homework is considered important and parents expect students to study each evening and this dominates family life (Cheng, 2011). Schools encourage parental involvement in their child's learning; home-room teachers visit their students' homes at least once a year and parents reinforce students' school progress at home (Barber, Donnelly and Rizvi, 2012). Because of the burden placed on students many local authorities, including Shanghai, have placed a limit on the number of hours of homework that schools can assign (Cheng, 2011). An estimate of over 80 per cent of children are sent to tutorial school for extra sessions/classes, with a strong focus on teaching students how to pass the examinations *(ibid)*. At an example school quoted in the press (Ellis and Bratu, 2011) students study 12 hours a day, including 3 hours of homework after dinner. There is homework allocated for every evening including weekends, and students study through the holidays to prepare for college entrance tests.

Shanghai participated in PISA for the first time in 2009 and was ranked first in the results for Maths, Science and Reading. Andreas Schleicher, OECD, considers Shanghai's performance on PISA particularly impressive in that there is a low level of variability between test scores from different schools (Ellis and Bratu, 2011). This could relate to the initiative of pairing a weak school with a stronger one to aid improvement *(ibid)*, as mentioned earlier. It is also noted that the performance in Shanghai is unlikely to be representative of China as a whole. In rural areas not all students have access to school and for those that do, the quality of school facilities is worse than in Shanghai *(ibid)*. Fewer rural children go on to college compared to Shanghai (Ellis and Bratu, 2011; Hays, 2012). David Barboza, writing in the New York Times in 2010 summed up education in Shanghai as follows:

> *The Shanghai students performed well, experts say, for the same reason students from other parts of Asia – including South Korea, Singapore and Hong Kong – do: Their education systems are steeped in discipline, rote learning and obsessive test preparation. .... But many educators say China's strength in education is also a weakness. The nation's education system is too test-oriented, schools here stifle creativity and parental pressures often deprive children of the joys of childhood, they say.* (Barboza, 2010)

Hong Kong

Because Hong Kong used to be under British rule, its education system has historically had a similar structure to that of the UK though affected by Chinese culture (OECD, 2011a). Hong Kong was returned to China in 1997 (after a 99 year lease ended) but maintains a considerable level of independence under the banner of 'one country, two systems' *(ibid)*.

There were very few public schools in Hong Kong until the 1950s, and until the 1970s/1980s some schools did not have their own premises but had to use vacant spaces (e.g. rooftops) (OECD, 2011a). Education only became compulsory in Hong Kong relatively recently; six years of primary education became compulsory from 1965 with the addition of three years of lower secondary in 1978 *(ibid)*. The Chinese belief that it is hard work and not family background that brings achievements is exemplified in Hong Kong, with high-performing students from varying backgrounds entering universities around the world *(ibid)*.

There was much dissatisfaction about schools in the 1990s (OECD, 2011a). There were excessive amounts of homework, much of which was regurgitation, and parents were generally unhappy with the education received and would send their children to international schools if they could afford to *(ibid)*. Teachers were dissatisfied with reported declining standards and motivation of learners, and employers were unhappy with the calibre of local graduating students *(ibid)*. Such dissatisfaction led to long term education reform which began in 1999 *(ibid)* and is due to be completed in 2016 when the first cohort finishes undergraduate study (Barber, Donnelly and Rizvi, 2012).

Interesting elements of this reform include a high level of consultation with the public about their ideas for what education should be like, as well as extensive consultation with, and involvement of, relevant experts (OECD, 2011a). The reform was informed by contemporary theories of learning *(ibid)*. The reform strategy and its implementation were carefully planned; a logical sequence for execution of elements of the strategy was ensured and all elements were linked to the overarching goal of improving teaching and learning (Barber, Donnelly and Rizvi, 2012). One of the major themes of the reform was to change the focus from 'teaching' to 'learning' and reduce the emphasis on memorisation of facts (OECD, 2011a).

As well as consulting the public in early stages of planning reforms, the Government went to considerable efforts to engage the public on changes to the curriculum introduced in 2009 and to develop teacher buy-in and ownership (OECD, 2011a; CIEB, 2013c). Day 'retreats' were held with senior representatives from 12 schools at a time, and later with groups of middle managers in order to prepare schools for implementation (OECD, 2011a). These sessions involved guest speakers talking about societal change (i.e. the need for educational change), and curriculum developers outlining curriculum reform and discussion of strategies for implementation *(ibid)*. Sessions began four years before implementation of the new curriculum *(ibid)*. In addition, briefings were held with the press in order to help engage the public with the changes *(ibid)*. This constitutes an unprecedented campaign to involve a large number of citizens and professionals in public discussion about educational goals (CIEB, 2013c). The result has been a widespread commitment to the new curriculum and education agenda *(ibid)*. This may have brought some challenges for teachers as they have been responsible for finding ways to implement the reform in schools, but it has also pushed them into exercising professional autonomy to adapt the reforms to their students (OECD, 2011a).

The Hong Kong Diploma of Secondary Education (HKDSE), which was introduced in 2009, is taken (usually) at age 17 years. It is interesting that the top available grade (5**) is worth more UCAS points than an A* at A level (145 and 140 points respectively) (Ma, 2013).

Hong Kong has attempted to move away from an exams-centred education system and from rote learning, and to embrace skills such as critical thinking and creativity, and provide well-rounded students (Hong Kong Higher Education, 2007). There is greater emphasis on group projects and open-ended assignments (Wikipedia, 2013a). There have also been efforts to gradually make more use of technology in the classroom in order to make learning more interactive (Li, 2012). However, the education reform has not eliminated the focus on 'quantity' of education or eliminated examination culture, with testing seen as a necessary element of education (Hong Kong Higher Education, 2007). Students in Hong Kong do well in international tests, but evidence suggests that their interest in learning and confidence that they can learn are relatively low (Li, 2012). Hong Kong primary students were ranked first in reading by PIRLS but close to the bottom for reading satisfaction and interest (Chan, 2012). This has been attributed to Confucianism by some, in that students are expected to work hard to meet their parents' expectations, but this does not necessarily convert to interest in learning (see Li, 2013).

Parents have high expectations of their children (Barber, Donnelly and Rizvi, 2012). Taking extra tutorial sessions is common. Statistics suggest a third of secondary school students had private tutoring in 2004–2005, (Wikipedia, 2013b), whilst current quotes suggest that 70 per cent of secondary school students have private tuition, (Adamson, 2013). Cram schools or tutorial schools attract many students for classes on exam technique, practising responses and tips on topics that may come up (Wikipedia, 2013a, 2013b). Some of the tutors have become very popular with celebrity status. These 'star tutors' or 'tutor kings/queens' appear on billboards and advertising and some have their own stylists and photographers to increase their popularity (The Hong Kong Standard, 2012; Wikipedia, 2013b). Such tutors may receive much higher salaries than secondary school teachers but may not have trained as teachers. (For more on the celebrity status of some tutors, see The Hong Kong Standard, 2012.)

Enrolment on study camps during school holidays, and/or asking teachers for additional texts that students can study during breaks is not uncommon (Mao, 2008). These kinds of activities are perhaps motivated by a strong desire by parents for their children to do well and not to need to work the long hours that they do when they enter employment – it is not uncommon for adults to work 70 or 80 hours a week *(ibid)*. (The average is 49 hours a week, according to a report by UBS, see Wikipedia, 2013c, and Hong Kong had the fifth longest yearly working hours of the countries studied.) The pressure to 'get ahead' and high parental expectations appear to foster a culture of test results mattering more than underpinning understanding (Mao, 2008). Competition between students for job and advanced school placements is fierce with access based on rankings in tests (Wikipedia, 2013a).

Hong Kong has invested resources in upgrading the quality of its teaching professionals (OECD, 2011a; Wikipedia, 2013a). Various undergraduate and postgraduate programmes are available for in-service and pre-service teachers (Wikipedia, 2013a). From 2002/03 all principals were required to undertake 50 hours of continuing professional development (CPD) a year, and from 2004/05 aspiring principals have to attain 'Certification for Principalship' *(ibid)*.

An interesting feature of the majority of Hong Kong schools is their strict codes of discipline and 'Demerit Points Systems' through which disciplinary offences are recorded (Wikipedia, 2013a). The record is included on students' report cards. Points can lead to suspension, expulsion and ultimately jeopardise whether students graduate and thus their future prospects (Wikipedia, 2013a).

The education system in Hong Kong has been described as 'spoon-fed' (Wikipedia, 2013a). Recent reforms are attempting to address this but whether teacher practices have changed substantially so that this is really no longer the case is difficult to tell from the available information. There are also some concerns that there are significant differences between schools in the quality of implementation and in performance (CIEB, 2013c).

Singapore

Singapore has developed within 50 years from a poor island, with high illiteracy and few natural resources (except for its deep water port), to a thriving economy with a high standard of living (CIEB, 2013d). Singapore became independent from Britain in 1958 (Spar, 2009), then briefly part of Malaysia from 1963 to 1965, when it was expelled and became independent again (Stewart, 2011; Wikipedia, 2013d). On independence, Singapore was potentially politically and economically vulnerable (Stewart, 2011). This led to a certain sense of urgency over policy making *(ibid)*. Politically Singapore has been very stable. For example, it had the same Prime Minister, Lee Kuan Yew, from 1959 to 1990 (Spar, 2009). Lee set ambitious economic goals and knew that education would be an important part of achieving them (CIEB, 2013d). Policies were informed by global benchmarking, were designed to support existing related policies and were carefully executed (CIEB, 2013d). There is a high degree of Government intervention in societal affairs. For example, enforced savings (Spar, 2009) and cases of demanding use of proper English rather than 'Singlish' in TV shows (see Srilal, 1999). However, generally the Government seem to be popular and respected as: goals are well communicated and largely well met, mistakes are admitted/resolved and the Government listens to views and responds via a complaints department/ministry (Spar, 2009). The Government promoted Confucian values (e.g. hard work and obedience to authority) and a culture of a Singaporean being someone who works hard and contributes to national economic success *(ibid)*. Singapore is openly described as a meritocracy in which high educational achievements, and test results, lead to advancement and good positions, and thus there is an intense focus on education and exams (Spar, 2009; Chia and Toh, 2012). Parents have high hopes for their children (e.g. aspirations for them to become Government administrators) and thus education is very important to Singaporean families (Spar, 2009). Parental involvement is seen as integral to students' success according to the education strategy in Singapore (Barber, Donnelly and Rizvi, 2012).

From Confucian values, there is a culture of it not being a case of intelligence, but of discipline and studying hard to do well. Confucianism also promotes respect for authority, which for students will relate to respect for their elders (Spar, 2009). The motivation to succeed generally comes from the home environment and parents instil in children the need to do well in school (Larkin, 2012). There is intensive parental involvement in their child's education *(ibid)*. Arguably, education has been the key to success for economic growth in Singapore, and this has been facilitated by Confucian values *(ibid)*.

The Government in Singapore is highly efficient with a focus on strategic planning and detailed execution (Stewart, 2011). Their policy development and implementation has been characterised as 'Dream, Design, Deliver' *(ibid)*. At the point of independence, Singapore had multiple religious groups, ethnic groups and no common language *(ibid)*. Lee believed in building a national identity and achieved this through various strategies such as mixing ethnic groups in Government-built housing and determining four official national languages *(ibid)*. Additionally, schools promote values of unity and national pride *(ibid)*. In a country of few natural resources, its people are seen as the main resource and as providing the capital for economic growth *(ibid)*.

Singapore's education system has developed in a number of phases. These have been described (Boon and Gopinathan, 2006; Stewart, 2011) as:

- **Survival-driven phase: 1959 to 1978.** At the time of becoming independent most Singaporeans were illiterate and unskilled, with only the rich having their children educated. Thus, this initial phase focused on increasing the number of schools, teachers and access as quickly as possible to provide basic education. Primary education was universal by 1965 and lower secondary by the early 1970s. However the quality of education was not high and there were high dropout rates.
- **Efficiency-driven phase: 1979 to 1996.** In 1979 the education system was revised into multiple pathways to reduce drop out, improve education quality, and develop a more technically-skilled labour force (to meet new economic goals to become a capital and skill-intensive country).
- **Ability-based, aspiration-driven phase: 1997 to present.** With the world economy changing and emphasis shifting to innovation, creativity and research, there were shifts in the education system to address this. A new educational vision was set out, 'Thinking Schools, Learning Nation', encompassing a range of initiatives including a focus on students developing creative thinking skills, a broadening of subjects, changes in school management and developing career paths and incentives for teachers.

Singapore has emphasised raising the quality of its teachers during the development of its education system (CIEB, 2013d; Pearson, 2013a).

Concerted efforts were made to raise the image of teaching and to provide training and better working conditions in the mid-1990s (Lim, 2012). Those entering teacher training are from the top third of their secondary school class (CIEB, 2013d). They receive training at a top HEI, where providing the appropriate values is emphasised as well as the appropriate knowledge and teaching skills *(ibid)*. There is a concerted effort to recruit and nurture talent for teaching *(ibid)*, including efforts to recruit mid-career individuals from industry (Pearson, 2012). Teachers are entitled to 100 hours of CPD (courses, conferences etc.) each year (Stewart, 2011; Pearson, 2012). Teacher pay is benchmarked against jobs in industry and schools have the flexibility to reward higher performing individual teachers or teams, there are several defined career tracks for teachers (e.g. leadership track, progression into the Ministry of Education, becoming a master teacher, becoming a senior specialist teacher in a particular area) and there are national teacher awards (Stewart, 2011; Pearson, 2012). The Singapore Government is mindful of how certain students may require additional support (e.g. students in single-parent households, ethnic minorities) and thus set up local community councils to identify and support families in need (CIEB, 2013e).

The school system in Singapore is sometimes accused of being too focused on grades and test performance and hence, high stress, with many students taking enrichment classes or additional tuition (Lim, 2012). Some now feel that the culture has become too pressurised with students having little 'down time' between school, extra-curricular activities, and outside study, and that this makes students unhappy (Koh, 2012). This is viewed by some as an undesirable but unavoidable result of success being viewed in quantitative terms *(ibid)*. The exams culture is promulgated by the Primary School Leaving Examination (PSLE). The PSLE is considered very important by parents, with some parents even taking a career break to support their child's preparation (Chua, 2012). PSLE results determine which school students can access and which stream students are assigned to (Express, Normal Academic, or Normal Technical). Some feel the PSLE and resultant streaming cause stress for parents and children (Yam, 2010; Chia and Toh, 2012), that the streaming is conducted at too young an age, and that it creates a climate of elitism (Yam, 2010). There is a current debate about whether the PSLE should be scrapped or changed in order to reduce the pressure it places on students and the strong focus on exams (Chia and Toh, 2012).

A lack of flair, creativity and individuality amongst students has also been a criticism of the Singapore education system (Watson, 2012). The new vision for education 'Thinking Schools, Learning Nation' was intended to encourage active learning and critical thinking in schools (Barber, Donnelly and Rizvi, 2012). In 2004, an initiative, 'Teach Less, Learn More', was developed to try to reduce rote learning and repetitive tasks and move towards more problem-based and deeper learning (CIEB, 2013d). "Holistic education" has also become a focus with a push to move towards cultivating creativity and innovation and being able to process information, as well as developing content knowledge (Lim 2012). Teachers are being given more leeway to find different approaches to learning and different ways to teach the syllabus *(ibid)*. Education Minister Heng Swee Keat is quoted (by Watson, 2012) to have said the future is:

> *… less about content knowledge, as content will have to be re-learnt and even un-learnt during one's lifetime….It is more about how to process information, discern truths from untruths, connect seemingly disparate dots and create knowledge even as the context changes. It is about developing an enduring core of competencies, values and character to anchor our young and ensure they have the resilience to succeed.*

The efforts to promote 21st century skills and competencies and to value a more holistic development of children through such initiatives appear to have had some impact (CIEB, 2013d). However, there may be a tendency for activities promoting the above to be secondary, or an 'add on' to covering the academic curriculum. Exam preparation and a focus on the approved textbooks continue to dominate. Stewart (2011) notes that some of the factors in Singapore's success are:

- integrated and forward-looking planning;
- close links between policy, research and educators;
- policies that align with each other making implementation more successful;
- small scale;
- a commitment to equity and merit;
- focus on Maths, Science and technical skills;
- strategic selection and development of teachers;
- continuous improvement.

Victoria, Australia

In 1788, Australia was established as a British penal colony and settlement by Europeans began (CIEB, 2013f). Immigration by Europeans was actively encouraged for many years as Australia, with its valuable natural resources but sparse population, felt vulnerable to their neighbours *(ibid)*. Wider immigration is now encouraged, though 92 per cent of the Australian population are of European descent *(ibid)*.

The teaching profession was demoralised by the depression of the 1890s with school buildings deteriorating and the curriculum in need of modernisation (Wikipedia, 2013e). In 1902 a Director of Education, Frank Tate, was appointed for Victoria and instigated modernisation, bringing in a child-centred pedagogy and broadening the subjects taught in primary school *(ibid)*. Until after the end of World War II Australia traded mostly with Britain, but broadened their horizons later, including a much increased focus on trading with Asia (CIEB, 2013f). The long-standing focus on key countries in the West has made Australians keen on benchmarking, and thus they sought insights on how leading countries had encouraged economic growth in determining how to move forward themselves *(ibid)*. This process identified that it would be important to their success to invest heavily in education and training to raise their standards *(ibid)*. Academic and vocational standards were created, the curriculum was developed and efforts were made to strengthen the teaching profession *(ibid)*.

Post World War II baby boom and immigration increased the school population in Victoria (Wikipedia, 2013e). Increasing class sizes resulted in the use of church halls and other temporary locations whilst new school premises were built *(ibid)*. There were also teacher shortages which led to married women returning to work *(ibid)*. By the 1970s Victoria's primary schools were evolving through child-centred pedagogy, school-based curriculum development, multiculturalism and genuine partnerships with parents *(ibid)*.

To improve teacher quality as part of recent educational reforms, the Australian Institute for Teaching and School Leadership (AITSL) was set up in 2010 (Pearson, 2013b). They co-ordinated the establishment of agreed national standards and expectations for teachers. This was a challenging task given the diversity of backgrounds of different states/territories but was achieved by involving 6000 practising teachers and principals to ensure teacher engagement and buy in *(ibid)*. The standards are currently being embedded. Teacher training is also being redeveloped to improve its quality and consistency *(ibid)*. A national requirement for a two-year postgraduate qualification as a prerequisite for teaching has been put in place and additionally those entering teacher training must be in the top 30 per cent of the population for numeracy and literacy skills *(ibid)*.

Education policy in Australia has been largely determined by the individual states and territories (CIEB, 2013f). However a National Programme of Assessment (NAP) began in 2008 and a National Curriculum was adopted from 2011 *(ibid)*. Victoria has merged their existing curriculum into the new National Curriculum in order to retain the former's local character.

Whilst Australia performs well in international comparisons of achievement, there is a wide gap between educational achievements of those from privileged and disadvantaged backgrounds (Pearson, 2013b). Australia has put considerable financial resources into its education system in recent years and instigated a number of initiatives to help weaker schools (CIEB, 2013f). For example, one initiative focuses on increasing students' information and communication technology (ICT) skills, and another on improving teacher training and retention through improved opportunities for professional development *(ibid)*.

Indigenous children tend to perform less well than others on achievement tests. Whilst it has been hypothesised that this might be due to disillusionment with school, a 1995 survey found no difference in attitudes to school between indigenous and other groups (Marks, 1998). Some sources feel that teachers sometimes assume that aboriginal students are likely to do less well and have low expectations of them rather than promoting an attitude that such students have just as much potential as anyone else (see for example, Korff, 2012). In 2009 the State Government of Victoria published a strategy titled 'Education for Global and Multicultural Citizenship' (Department of Education and Early Childhood Development [DEECD], 2009) to support schools by providing a renewed vision for global and multicultural education. It emphasises themes such as improving educational outcomes for all students, promoting social cohesion and enhancing the engagement, wellbeing and sense of belonging for all students.

A report commissioned by the Prime Minister's Science, Engineering and Innovation Council (PMSEIC, 2009) suggests that Australia has an attitude problem with regard to learning. Anti-social behaviour is reportedly becoming more evident both inside and outside of schools, with increasing numbers of children who are inattentive, disruptive, disengaged and under-motivated (Australian Primary Principals Association [APPA], 2008). There is a lack of enthusiasm for learning and the traditional cultural attitude to learning in Australia is a relaxed 'he/she'll be alright' approach (Milburn, 2010). There are examples of how a proactive head teacher can turn around a school through the 'no-excuses model', which opposes the view that socio-economic status and innate ability are the main determinants of success, and places greater emphasis on the quality of teachers *(ibid)*. This may include focussing on encouraging different teaching methods, raising expectations, and raising student attainment which in turn improves parents' attitudes to education by seeing their child's improvement *(ibid)*. This perspective advocates teachers placing high expectations on themselves as well as on their students. However, such practices may be relatively rare, and teachers do not feel that their own efforts to improve their teaching would be recognised *(ibid)*.

The PMSEIC report calls for a cultural shift in attitude with society needing to value education (Milburn, 2010). A campaign to lift the status of teaching is recommended, and another to make it part of the Australian identity to value education *(ibid)*. Some argue that it will take specialised support and additional resources to reverse student underperformance and family attitudes towards learning in schools with high numbers of disengaged students and parents *(ibid)*.

New Zealand

Primary education became compulsory in New Zealand in 1877 when an Education Act was passed (Wikipedia, 2013f). In 1900 there were fewer than 10 per cent of students who went on to secondary school, for which there were fees *(ibid)*. As the 20th century progressed there was increased need for skilled tradespeople and administrators and the secondary sector expanded *(ibid)*. In 1914 another Education Act was passed and secondary schools were required to offer free education to all who could pass a proficiency exam *(ibid)*. By 1917 the percentage of students attending secondary school had increased to 37 per cent *(ibid)*. The schools were similar to Grammar Schools in England with a traditional curriculum suitable for those intending to go to university *(ibid)*. Technical Schools were introduced with the intention of offering a 'relevant' alternative of equal status *(ibid)*. However, these tended to have a stigma attached to them and the balance of attendees at the two school types tended to run along class lines *(ibid)*. The Thomas Report of 1944 resulted in a common, core curriculum drawing on academic and practical areas, aimed to be appropriate for students of varying ability, interests and background. Schools resisted this change by streaming students by ability *(ibid)*.

Although New Zealand once had one of the strongest economies in the world (at the end of the 19th century its raw materials and privileged trade relationship with Britain were to its advantage), by the 1980s New Zealand was in financial trouble with rising inflation and unemployment, partly due to dramatic increases in oil prices and heavy borrowing from abroad for major projects (CIEB, 2013g; Wikipedia, 2013g). The national education system had a good reputation both internally and externally, but was nonetheless reformed as part of the then Labour Government policies to address the economic situation (CIEB, 2013g). In 1989, the reportedly bureaucratic and out of touch Department of Education and its regional offices were replaced with a much smaller Ministry of Education *(ibid)*. The ethos driving this was to move governance towards the school level, with each school having its own board *(ibid)*. The National Party replaced Labour in 1990 and emphasis turned towards using market mechanisms to make education effective and to good management being key to school success *(ibid)*. However, Labour's 'Tomorrow's Schools' are still the basis of today's New Zealand schools *(ibid)*. Devolving responsibility to the local level strengthened some schools (usually those serving financially better off and better educated parents), substantially weakened others (usually those serving communities with less well educated parents, where the school boards found it harder to recruit good teachers), and for others made little difference *(ibid)*. Whilst 'Tomorrow's Schools' have widened the gap between weaker and stronger achievement, overall achievement levels remain fairly stable, and performance on international comparisons, such as PISA, is high *(ibid)*.

Of OECD countries, New Zealand has a relatively equal income distribution with less poverty than elsewhere (CIEB, 2013g). Additionally, New Zealand has a long history of national welfare provisions. For example, it is 8th amongst 39 countries in percentage of Gross Domestic Product (GDP) devoted to family benefits *(ibid)*. New Zealand has a relatively high number of 3 to 6 year olds in early childhood education and day care with relatively more of this provision funded by the Government *(ibid)*. Thus, children in New Zealand may be starting primary school with a developmental advantage over peers in other developed countries *(ibid)*.

In the early 1990s the New Zealand Qualifications Authority (NZQA) was set up as a separate body to the Ministry and with authority to create a comprehensive Qualifications Framework ranging from high school leaving qualifications to doctorates and including vocational qualifications (CIEB, 2013g). This harmonised framework has the advantage that HEIs and employers recognise these qualifications and students know that their achievements will be recognised *(ibid)*. This strong system of qualifications is thought to have a positive influence on the quality of teaching in New Zealand schools and incentivise students to work hard *(ibid)*. The main qualifications taken by secondary school students is the National Certificate of Educational Achievement (NCEA) which is available at three levels. Some schools offer IGCSEs, International AS and A levels through Cambridge International Examinations, though these are not registered on New Zealand's National Qualifications Framework (Wikipedia, 2013h).

New Zealand also has strengths regarding reading instruction. A system known as 'Reading Recovery' is used which helps teachers to identify children struggling to read and write, and provides teachers with skills in tutoring techniques to help these children (CIEB, 2013g). This has reportedly been very successful with a positive influence on the teaching of reading beyond just those students in need of particular help *(ibid)*. Arguably, addressing literacy problems early can prevent knock-on problems for students across subjects and through their schooling so the effects of this system are thought to be far reaching *(ibid)*. In addition, Numeracy and Literacy Development Projects in 2000 and 2004 supported improvement in teachers' content knowledge in these areas *(ibid)*. Also, in 2008-2009, a programme began to update national standards for literacy and numeracy with accompanying national assessment *(ibid)*.

New Zealand culture is strongly influenced by British and European customs but there is also a strong influence from Māori and Polynesian traditions (Wikipedia, 2013i). Māori settlers arrived on the islands first (before 1300) but Europeans, arriving considerably later (around 200 years ago) had a dramatic effect *(ibid)*. The Treaty of Waitangi was signed in 1840 to facilitate peaceful relations between Māori and European New Zealanders (Pākehā) *(ibid)*. However the treaty was not initially effective and the New Zealand land wars broke out in 1845 *(ibid)*. This had a negative impact on the place of Māori culture, but its influence has been regained in recent decades (Wikipedia, 2013i). In order to promote understanding between Māori and Pākehā, biculturalism and the Treaty of Waitangi were made part of the school curriculum in the late 20th century *(ibid)*.

Pākehā culture has developed mostly from that of the original British settlers, but there are some distinct differences which have increased over time (Wikipedia, 2013i). An interesting feature is that Pākehā culture has a stronger emphasis on egalitarianism and anti-intellectualism (Keown, Parker and Tiakiwai, 2005) and the idea that most people can do most things if they put their minds to it (Wikipedia, 2013i). Intellectual activity is not particularly well regarded. The focus is instead on the 'kiwi ingenuity' of finding a practical 'what works' solution to a problem (Keown, Parker and Tiakiwai, 2005) rather than much emphasis being placed on applying a theory (Wikipedia, 2013i). Modesty is valued (Keown, Parker and Tiakiwai, 2005; Wikipedia, 2013i) which, on one hand, could potentially minimise any differential expectations of students and an attitude of equality in the classroom may be conducive to learning. On the other hand, valuing modesty and the distrust and dislike of those who boast of their own merits has been criticised as discouraging ambition and individual achievement (Wikipedia, 2013i). These are potentially conflicting influences on student attitudes to learning.

It is also interesting that one of the effects of European colonisation was that in the 1830s many Māori converted to Christianity and consequently learnt to read and write (Wikipedia, 2013i). In 1867 the Native Schools Act was passed (by a Pākehā dominated parliament) which required Māori children to be taught mostly in English *(ibid)*. Additionally, most Māori parents encouraged their children to learn the English language to be able to function socially and economically *(ibid)*. Whilst there was later dissatisfaction with this Eurocentrism and demand for equal recognition of Māori culture (which occurred in time) *(ibid)*, these historical factors may have influenced current high performance in literacy and the emphasis on literacy through the 'Reading Recovery' method.

Up to the 1980s, New Zealand was claimed to be a classless society with a small difference between the salaries of higher and lower paid workers (Wikipedia, 2013i). Economic reforms in the 1980s and 1990s changed this due to international capital, commerce and advertising. Cheap imports damaged local manufacturing and jobs were lost. The gap between the richest and poorest New Zealanders increased *(ibid)*. Some argue that ethnicity takes the place of class in New Zealand, as Māori and other Polynesians tend to earn less and have a lower standard of living and less education *(ibid)*. According to a report by Mahuika and Bishop (2010), Māori students are treated differently in mainstream schools, often negatively. Bishop, *et al*. (2003, cited in Mahuika and Bishop, 2010) found that teachers identified the main influences on the educational achievement of Māori children as the students themselves, their homes and/or the structure of the school, thus labelling lack of achievement in deficit terms and as not due to the classroom. Mahuika and Bishop argue that this influences the quality of teachers' relationships with Māori students and leads to teachers having low expectations of these children. Based on this study and other findings, Mahuika and Bishop argue that deficit theorising by teachers is the main reason for the lower educational success of Māori students. They suggest that changing these positionings would allow teachers to realise their ability to affect the situation and cause change. There is now more common, but not universal, acknowledgment that Māori learners may have their own specific needs that are not the same as those of their peers (Mahuika and Bishop, 2010). Finding successful ways to deal with the differences, if acknowledged, is also partial *(ibid)*. Mahuika and Bishop argue that teachers do not understand the role that culture plays in learning and educational assessment and do not know how to adjust for this in their teaching and assessment strategies. For example, traditionally Māori people learnt through direct experience in the natural world and understanding particular ideas from this holistic perspective rather than learning in a decontextualised way which is common in Western classrooms. Arguably taking into account the different learning and assessment styles relevant to the culture would be beneficial *(ibid)*.

Despite New Zealand doing well in international comparisons, the Ministry of Education is not standing still in terms of its ongoing educational aims. The New Zealand Ministry of Education published a recent report by the New Zealand Council for Educational Research focused on developing teaching and learning for the future (Bolstad *et al*., 2012). This document puts forward a strategy for developing 'future-oriented' or 21st century learning through 'unbundling' current school practices and rethinking them, drawing on existing examples of forward thinking practice, and focussing efforts around a number of key themes including:

- personalised learning;
- new views of equity, diversity and inclusivity;
- a curriculum that uses knowledge to develop learning capacity;
- "Changing the script" – rethinking learners' and teachers' roles;
- a culture of continuous learning for teachers and educational leaders;
- and new kinds of partnerships and relationships between the school and the community.

Summary of cultural and societal factors that may contribute to high performance

This section summarises some of the societal and culturally related factors that are common to some high-performing jurisdictions. These factors could be amongst the reasons for the success of their education systems, though it would be unwise to assume simple causality. Various previous reports by different authors have identified overlapping subsets of these factors (e.g. Reynolds and Farrell, 1996; Green, 1997; Elliott and Phuong-Mai, 2008; OECD, 2011b; Tucker 2011; Economist Intelligence Unit, 2012; CIEB, 2013h).

Status of teaching as a profession/quality of teaching/teacher professional development

In most high-performing jurisdictions teaching is a highly respected career. Thus, access to teaching courses can be selective, with only those with high achievement selected. Often, teacher pay is good (e.g. Singapore, South Korea) or at least stable and reliable (e.g. Shanghai) thus making it an attractive option. There is rigorous teacher training and requirements for continued professional development (e.g. 100 hour annual CPD entitlement and designated career paths in Singapore; group lesson planning and lesson observations in Shanghai). The apparent importance of teacher development/quality/status to the success of education systems in high-performing jurisdictions has previously been noted by the OECD's Andreas Schleicher (2011) and by Pearson's 'The Learning Curve' report (Economist Intelligence Unit, 2012). The latter argues that countries with the best performing teachers work to attract talented individuals and train them throughout their careers. They argue that the effect of good teaching goes beyond that of positive educational achievements, also influencing wider societal issues (e.g. lower levels of teenage pregnancy, greater tendency to save for retirement).

In a short summary of research findings on this theme, Canadian Education Association (CEA) (2011) concluded that four key factors that affect teaching quality, and thus learning, are:

- teacher knowledge of good/effective teaching and learning practices;
- a degree in the subject taught;
- teachers' verbal ability and literacy level;
- continuing professional development (which is necessary for improvements in teaching quality, particularly when there are changes such as to the curriculum or student population).

Cultural attitudes that highly value education, learning and hard work

In most high-performing jurisdictions, education is highly valued and the 'cornerstone' of their culture. For example, in Eastern countries such as China and Singapore, Confucian values are culturally embedded and emphasise that hard work, discipline and perseverance bring success. This is very much in contrast with Western views that 'you're either smart or you're not' (Elliott and Phuong-Mai, 2008; Yeung and Yeung, 2008; Barber, 2012). The historical context of the Civil Examination in China and folklore surrounding it has also influenced societal views on the value of education, leading to high aspirations, regardless of background. A side effect of respect for education in Chinese cultures is that there is often a supportive, pro-learning peer culture (Elliott and Phuong-Mai, 2008). The latter tends not to be the case in some Western cultures, where studying hard may lead to criticism or teasing from others (e.g. Elliott and Phuong-Mai, 2008). Finland is another example; here there is a long history of emphasis on literacy and reading skills and a societal commitment to high achievement (Andrews, 2010, cited in Oates, 2010). New Zealand is an interesting case as the anti-intellectual and egalitarian cultural perspective embodies both a view that most people can do most things if they apply themselves, and a negative view of anyone immodest or doing better than others.

Positive societal attitudes towards education have been identified as a possible factor by various researchers and commentators. Green (1997), for example, concluded that:

> …*the essential difference between the compulsory school systems of the high achieving countries as compared with the lower achieving countries would appear to be that the former have both a culture and certain institutional mechanisms which encourage high aspirations and achievement among a wide majority of children*.
> (p.122).

One piece of evidence that cultural views valuing education do affect performance is that Asian-American students tend to outperform their American peers even when ability and socio-economic status are taken into account (see Elliott and Phuong-Mai, 2008). Such students may be influenced more by the values of their parents and their own culture, than the (usually) anti-intellectual views of their class mates *(ibid)*.

Parental involvement in education/parental expectations

In some jurisdictions where students perform well in international tests, parental involvement or expectations may be a factor. For example, in Shanghai parents have high expectations of students and a strong desire for them to do well, expecting them to study each evening; parental involvement in reinforcing students' learning is encouraged by schools. Parents in Hong Kong and Singapore also have high expectations of their children and in all three of these jurisdictions many parents send their children to tutorial schools or other forms of extra tuition. Parents make sacrifices for their children's education and expect them to repay the sacrifice (Elliott and Phuong-Mai, 2008). In Finland, learning begins in the home and early literacy is an important element of society (Oates, 2010). In 1686 literacy was made a requirement for marriage in Finland and this may be part of the background to continued emphasis on learning in the home, with early literacy providing the vehicle for further learning. In Canada the degree to which parents engage their children in literacy activities is also high. Positive parental engagement in students' learning seems to be less the case in Australia, despite its good performance in international comparisons. However, examples of strategic efforts (at the school level) to improve student performance have shown that a 'virtuous cycle' can be created with improvements in student abilities leading to parents having more respect for education and encouraging their children more.

Respect for authority and for the knowledge of teachers and parents

There is a culture of respect for authority and one's elders in some high-performing jurisdictions, particularly those with Chinese heritage. Respect for the knowledge of their teachers may be a factor in student behaviour in school and the tendency to pay attention in class. Respect for parents may be a factor in engagement in homework and extra tuition, and willingness to work hard to try to meet their parents' expectations.

An exam-driven culture, but with moves towards 21st century learning

In some high-performing jurisdictions there has been a long history of education being focused on exam preparation and a view that only test results count (e.g. the history of the Civil Examination may be a strong cause of this in Chinese society). Whilst such culturally-embedded views may potentially place limits on the kind of learning taking place, they are part of a structure that reinforces the importance of studying in order to do well in the future, and a strong focus on exam technique may be one reason for the high performance of some jurisdictions in international comparative tests. In recent decades, policy makers in these locations have realised that students need more than rote-learned facts in order to compete in the global market and that skills in areas such as problem solving, application to real life situations, critical and creative thinking and teamwork are important. A number of initiatives have been instigated in Singapore, Hong Kong and Shanghai to encourage less didactic lessons with use of a wider variety of teaching methods and classroom practices, in order to facilitate the development of such skills.

Canada is a contrasting case, in that traditionally assessment has been classroom-based rather than exam-based. However, students have experience of provincial tests and take exams at age 18 years as part of their high school Diplomas.

Educating the 'whole student'

Tucker (2011) noted that high-performing systems focus on educating the whole student. In Canada there are cultural values around societal harmony being important, perhaps giving an educational focus beyond exams and assessment. A specific example comes from Ontario where there are high numbers of 'new Canadians' in need of particular support with language. As part of the 'student success' strategy, teachers meet to discuss students in need of support, or at risk of dropping out, and strategies for doing so are decided and then taken forward. Strong ethics and values are also important in Singapore and are seen as an important outcome of education in their own right. In Singapore there are also specific groups based in local areas to provide support for families whose students might do less well because of background. Moves towards wider teaching styles and broader goals for learning (e.g. 21st century skills) also represent moves towards educating the 'whole student' in Eastern countries.

Involvement of teachers in policy making/teacher and school autonomy

Some high-performing jurisdictions have made concerted efforts to involve, or at least consult, teachers and others in policy development and this may be a cultural or policy factor affecting the success of educational systems. Tucker (2011) has commented on this theme. Hong Kong is the strongest example of this with teachers and the public widely consulted on educational reforms. This helped with ensuring 'buy in' for the reforms and hence eased their implementation. Where education decisions are led by school level views in a 'bottom up' manner this may also have some advantages. The Netherlands and Hong Kong have high levels of school autonomy. China, whilst traditionally very hierarchical with a top-down education system, has made moves to devolve some power to the local level and Shanghai in particular has flexibility to pilot new ideas. In Canada, education decisions have traditionally been under the control of district boards, with the state having a supportive role; there is no country-wide Education Department. Whilst there may have been some degree of shift towards central organisation, the boards still have considerable control over how they do things. In Ontario, teachers were involved in designing the agenda for reform, and thus were committed to its implementation. However, whilst engaging teachers in policy and decentralism appear to be a feature of some high-performing jurisdictions, they are are not necessarily common to all.

Cohort size

It is notable that most of the highest performing jurisdictions in comparative national tests are considerably smaller than England. There were around 650,000 16 year olds in England in 2011 (Office for National Statistics, 2013), compared to about a tenth of that in some high-performing jurisdictions. (For example, around 60,000 in New Zealand [Education Counts, 2012], 46,000 in Alberta [Alberta Education, 2013b] and 73,000 in Hong Kong [HKEAA, 2012]). Alexander (2010) comments that at a simple level some of the best performing school systems appear to be small (and rich), though emphasises that it would be 'grossly simplistic' to assume cause and effect and eludes to the complexity of factors involved.

Spending

Whilst investing funds in education is important and can contribute to improvement, there are other factors and it is not as simple as 'more spending equals better results' or that 'paying teachers more equals better results'. Nisbet (2012) used UNESCO data to put a number of jurisdictions in order by percentage of GDP spent on education. The results did not give the same order as the rank order of jurisdictions by key international tests such as PISA and TIMSS. 'The Learning Curve' report (Economist Intelligence Unit, 2012) concludes that income does matter to education, but that the surrounding culture such as attitudes to learning may matter more and that cultural change in relation to education and ambition is needed to increase educational achievement outcomes.

Summary

This article has explored the cultural and societal context within which education takes place, with a focus on six jurisdictions whose students perform highly in international tests. A number of existing commentaries on the features of high-performing jurisdictions which may contribute to their success were particularly useful, along with a wider set of varied sources.

Governments are increasingly looking to the education systems of other countries to improve performance in their own, but the cultural and societal contexts of jurisdictions must be taken into account before any simplistic 'borrowing' of policy. The success of any education system will be a result of a complex interaction of different factors. This article has drawn together some of what is known about the cultural and societal factors potentially aiding success in high-performing education systems, though the interactions of different cultural and policy factors should not be forgotten in interpreting the list of factors above. In addition, the factors above are not universal to all high-performing jurisdictions and confirming that particular factors are indeed causes of success is difficult. Further to this, comparing features of weak and average performing jurisdictions would be needed to confirm whether factors discussed here are relatively unique to high-performing jurisdictions.

References

Adamson, B. (2013) 'Hong Kong'. In McClatchey, C., What is the key to a successful education system? *BBC News*, 7 February 2013. Available online at: http://www.bbc.co.uk/news/uk-21354932 (Accessed 13 December 2013).

Alberta Education (2013a) *Diploma Examinations*. Edmonton: Alberta Education. Available online at: http://education.alberta.ca/admin/testing/diplomaexams.aspx (Accessed 13 December 2013).

Alberta Education (2013b) *Student Population by Grade and School Year*. Edmonton: Alberta Education. Available online at: http://education.alberta.ca/apps/statistics/studpop/studpop.asp?grade=10to12 (Accessed 13 December 2013).

Alexander, R.J. (2010) "World class schools" – noble aspiration or globalised hokum? *Compare: A Journal of Comparative Education*, **40**(6), 801–817.

Anonymous (2011) The Disadvantage of Diploma Exams for Alberta Students. *'Have your say' section of Alberta Education website*, 14 December 2011. Available online at: http://ideas.education.alberta.ca/engage/idea-archive/general-idea-archive/the-disadvantage-of-diploma-exams-for-alberta-students/ (Accessed 13 December 2013).

Australian Primary Principals Association (APPA) (2008). *Position paper*. Kingston: APPA. Available online at: http://www.appa.asn.au/papers/Current-educational-issues.pdf (Accessed 13 December 2013).

Barber, M. (2012) Lessons on education from Singapore. *The Guardian*, 22 August 2012. Available online at: http://www.guardian.co.uk/commentisfree/2012/aug/22/lessons-education-singapore-gove (Accessed 13 December 2013).

Barber, M., Donnelly, K. & Rizvi, S. (2012) *Oceans of innovation: The Atlantic, the Pacific, global leadership and the future of education*. London: Institute for Public Policy Research.

Barboza, D. (2010) Shanghai schools' approach pushes students to top of tests. *New York Times*, 29 December 2010. Available online at: http://www.nytimes.com/2010/12/30/world/asia/30shanghai.html?ref=davidbarboza (Accessed 13 December 2013).

Bolstad, R., Gilbert, J., McDowall, S., Bull, A., Boyd, S. & Hipkins, R. (2012) *Supporting future-oriented learning and teaching – a New Zealand perspective*. Wellington: New Zealand Ministry of Education. Available online at: http://www.educationcounts.govt.nz/__data/assets/pdf_file/0003/109317/994_Future-oriented-07062012.pdf (Accessed 13 December 2013).

Boon, C.G. & Gopinathan, S. (2006) *The Development of Education in Singapore since 1965*. Background paper prepared for the Asia Education Study Tour for African Policy Makers, June 18–30 2006. Available online at: http://siteresources.worldbank.org/EDUCATION/Resources/278200-1121703274255/1439264-1153425508901/Development_Edu_Singapore_draft.pdf (Accessed 13 December 2013).

Bradshaw, J. (2012) Parents turn Canadian students into world-class readers. *The Globe and Mail*, 11 December 2012. Available online at: http://www.theglobeandmail.com/news/national/education/parents-turn-canadian-students-into-world-class-readers/article6187697/ (Accessed 13 December 2013).

Canadian Education Association (CEA) (2011) *The Facts on Education: Does Teaching Quality Make a Difference?* Toronto: Canadian Education Association and Ontario Institute for Studies in Education, University of Toronto. Available online at: http://www.cea-ace.ca/sites/cea-ace.ca/files/cea-2011-foe-teaching-quality.pdf (Accessed 13 December 2013).

Canadian Education Association (CEA) (2012) *Public Survey Results*. Toronto: CEA. Available online at: http://www.cea-ace.ca/sites/cea-ace.ca/files/cea-2012-public-survey-results.pdf (Accessed 13 December 2013).

Center on International Education Benchmarking (CIEB) (2013a) *Canada Overview*. Washington: CIEB. Available online at: http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/canada-overview/ (Accessed 13 December 2013).

Center on International Education Benchmarking (CIEB) (2013b) *Shanghai-China Overview*. Washington: CIEB. Available online at: http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/shanghai-china/ (Accessed 13 December 2013).

Center on International Education Benchmarking (CIEB) (2013c) *Hong Kong Overview*. Washington: CIEB. Available online at: http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/hong-kong-overview/ (Accessed 13 December 2013).

Center on International Education Benchmarking (CIEB) (2013d) *Singapore Overview*. Washington: CIEB. Available online at: http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/singapore-overview/ (Accessed 13 December 2013).

Center on International Education Benchmarking (CIEB) (2013e) *Singapore: Education for All, Student Support Systems*. Washington: CIEB. Available online at: http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/singapore-overview/4136-2/ (Accessed 13 December 2013).

Center on International Education Benchmarking (CIEB) (2013f) *Australia Overview*. Washington: CIEB. Available online at: http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/australia-overview/ (Accessed 13 December 2013).

Center on International Education Benchmarking (CIEB) (2013g) *New Zealand Overview*. Washington: CIEB. Available online at: http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/new-zealand-overview/ (Accessed 13 December 2013).

Center on International Education Benchmarking (CIEB) (2013h) *Top Performing Countries*. Washington: CIEB. Available online at: http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/ (Accessed 13 December 2013).

Chan, C. (2012) Pupils make champion readers – for exams only. *The Hong Kong Standard*, December 13 2012. Available online at: http://www.thestandard.com.hk/news_detail.asp?pp_cat=11&art_id=129254&sid=38464573&con_type=1&d_str=20121213&isSearch=1&sear_year=2012 (Accessed 13 December 2013).

Cheng, K. (2011) 'Shanghai: How a big city in a developing country leaped to the head of the class'. In Tucker, M.S. (Ed.) *Surpassing Shanghai: An agenda for American education built on the World's leading systems*. Cambridge, Mass.: Harvard Education Press. [Also published as part of: OECD (2011) 'Shanghai and Hong Kong: Two distinct examples of education reform in China'. In *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education*, OECD Publishing. Available online at: http://dx.doi.org/10.1787/9789264096660-en] (Accessed 13 December 2013).

Chia, S. & Toh, K. (2012) Scrapping PSLE not the solution Lawrence Wong. *The Straits Times*, September 22 2012. Available online at: http://www.straitstimes.com/breaking-news/singapore/story/scrapping-psle-not-the-solution-lawrence-wong-20120922 (Accessed 13 December 2013).

Chua, C.H. (2012) Mum, thanks for quitting job for PSLE. *The Straits Times*, 30 October 2012. Available online at: http://www.straitstimes.com/premium/forum-letters/story/mum-thanks-quitting-job-psle-20121030 (Accessed 13 December 2013).

Department of Education and Early Childhood Development (DEECD) (2009) *Education for Global and Multicultural Citizenship: A Strategy for Victorian Government Schools 2009-2013*. Melbourne: DEECD, State Government of Victoria. Available online at: http://www.eduweb.vic.gov.au/edulibrary/public/commrel/policy/multicultural-ed-strategy.pdf (Accessed 13 December 2013).

Economist Intelligence Unit (2012) *The Learning Curve: Lessons in country performance in education 2012 report*. Pearson. Available online at: http://thelearningcurve.pearson.com/the-report (Accessed 13 December 2013).

Education Counts (2012) *Student Numbers*. Wellington: New Zealand Ministry of Education. Available online at: http://www.educationcounts.govt.nz/statistics/schooling/july_school_roll_returns/6028 (Accessed 13 December 2013).

Ell, J.F. (2002) *History of Public Education*. The Alberta Teachers' Association. Available online: http://www.teachers.ab.ca/Teaching%20in%20Alberta/History%20of%20Public%20Education/Pages/Index.aspx (Accessed 13 December 2013).

Elliott, J.G. & Phuong-Mai, N. (2008) 'Western influences on the East, Eastern influences on the West: Lessons for the East and West.' In Tan, O.S., McInerney, D.M., Liem, A.D. & Tan, A. (Eds.) *What the West can learn from the East: Asian perspectives on the psychology of learning and motivation*. Charlotte, North Carolina: Information Age.

Ellis, R. & Bratu, B. (2011) How Shanghai's students stunned the world. *NBC Nightly News*, 2 November 2011. Available online at: http://www.msnbc.msn.com/id/44642475/ns/nbcnightlynews/t/how-shanghais-students-stunned-world/ (Accessed 13 December 2013).

Flanagan, T. (2011) Alberta's education system offers lesson in competition. *The Globe and Mail*, 11 October 2011. Available online at: http://www.theglobeandmail.com/commentary/albertas-education-system-offers-lesson-in-competition/article4403440/ (Accessed 13 December 2013).

Gaffield, C. (2012) 'History of Education'. In *The Canadian Encyclopedia*. Historica-Dominion Institute. Available online at: http://www.thecanadianencyclopedia.com/articles/history-of-education (Accessed 13 December 2013).

Green, A. (1997). *Education, Globalization and the Nation State*. London: Macmillan. Available online at: http://www.tandfonline.com/doi/pdf/10.1080/0142569980190407 (Accessed 16 December 2013).

Hargreaves, A. (2011) Canada's culture of excellence in education. *The Toronto Star*, 26 January 2011. Available online at: http://www.thestar.com/opinion/editorialopinion/article/928064--canada-s-culture-of-excellence-in-education (Accessed 13 December 2013).

Hays, J. (2012) *Chinese Education System*. Facts and Details website. Available online at: http://factsanddetails.com/china/cat13/sub82/item1718.html (Accessed 13 December 2013).

HKEAA (2012) 2012 *Hong Kong Diploma of Secondary Education (HKSDE) Examination Results Released*. Hong Kong: Hong Kong Examinations and Assessment Authority. Available online at: http://www.hkeaa.edu.hk/DocLibrary/Media/PR/20120719_HKDSE_Results_ENG_FULL.pdf (Accessed 13 December 2013).

Hong Kong Higher Education (2007) *Hong Kong's Education System*. Hong Kong: Hong Kong Higher Education. Available online at: http://studyinhongkong.edu.hk/eng/01hkesystem.jsp (Accessed 13 December 2013).

The Hong Kong Standard (2012) *Hong Kong's celebrity tutors*, 16 October 2012. Available online at: http://www.thestandard.com.hk/news_detail.asp?we_cat=16&art_id=127333&sid=37917717&con_type=1&d_str=&fc=4 (Accessed 13 December 2013).

International Qualifications Assessment Service (IQAS) (2007). *International Education Guide for the assessment of education from China*. Edmonton, Canada: International Qualifications Assessment Service, Government of Alberta. Available online at: http://eae.alberta.ca/documents/WIA/WIA-IM_iqas_china_profile.pdf (Accessed 13 December 2013).

Keown, P., Parker, L. & Tiakiwai, S. (2005) *Values in the New Zealand Curriculum*. Wellington, New Zealand: Ministry of Education. Available online at: http://www.waikato.ac.nz/wmier/research/projects/values-in-the-new-zealand-curriculum (Accessed 13 December 2013).

Koh, M. (2012) Why parents push kids so hard for PSLE. *The Straits Times*, 1 October 2012. Available online at: http://www.straitstimes.com/premium/forum-letters/story/why-parents-push-kids-so-hard-psle-20121001 (Accessed 13 December 2013).

Korff, J. (2012) *Common mistakes when teaching Aboriginal studies*. Creative Spirits website. Available online at: http://www.creativespirits.info/aboriginalculture/education/common-mistakes-when-teaching-aboriginal-studies (Accessed 13 December 2013).

Larkin, B. (2012) Education dominates Singapore's different culture. *The Plain Dealer*, 3 November 2012. Available online at: http://www.cleveland.com/opinion/index.ssf/2012/11/education_dominates_in_singapo.html (Accessed 13 December 2013).

Li, J. (2012) Evolution of learning. *The Hong Kong Standard*, 11 December 2012. Available online at: http://www.thestandard.com.hk/news_detail.asp?we_cat=16&art_id=129161&sid=38449985&con_type=1&d_str=20121211&fc=4 (Accessed 13 December 2013).

Li, J. (2013) Beyond the grades. *The Hong Kong Standard*, 15 January 2013. Available online at: http://www.thestandard.com.hk/news_detail.asp?we_cat=16&art_id=130084&sid=38696570&con_type=1&d_str=20130115&fc=4 (Accessed 13 December 2013).

Lim, R. (2012) Singapore wants creativity not cramming. *BBC News*, 23 May 2012. Available online at: http://www.bbc.co.uk/news/business-17891211 (Accessed 13 December 2013).

Ma, M. (2013) Education a brighter spot. *The Hong Kong Standard*, 3 January 2013. Available online at: http://www.thestandard.com.hk/news_detail.asp?pp_cat=21&art_id=129709&sid=38603021&con_type=1&d_str=20130103&isSearch=1&sear_year=2013 (Accessed 13 December 2013).

Mahuika, R. & Bishop, R. (2010) *REVIEW PAPER 6: Issues of culture and assessment in New Zealand education pertaining to Māori Students*. University of Waikato: School of Education. Available online at: http://assessment.tki.org.nz/Media/Files/Mahuika-R.-and-Bishop-R.-Issues-of-culture-and-assessment-in-New-Zealand-education-pertaining-to-Maori-students-University-of-Waikato (Accessed 13 December 2013).

Mao, F. (2008) *Beating the Unbeatable Foe: An educationally (for Hong Kong) incorrect post*. Available online at: http://faimao.blogspot.co.uk/2008/06/beating-unbeatable-foe-educationally.html (Accessed 13 December 2013).

Marks, G. (1998) *Attitudes to school life: their influences and their effects on achievement and leaving school*. Camberwell, Victoria: Australian Council for Educational Research. Available online at: http://research.acer.edu.au/lsay_research/62 (Accessed 13 December 2013).

Matsumoto, L. (2002) A Brief History: The Education Reform Movement in Alberta. *ATA Magazine*, **83**, 1. Available online at: http://www.teachers.ab.ca/Publications/ATA%20Magazine/Volume%2083/Number%201/Articles/Pages/A%20Brief%20History%20The%20Education%20Reform%20Movement%20in%20Alberta.aspx (Accessed 13 December 2013).

Mehta, J.D. & Schwartz, R.B. (2011) 'Canada: Looks a lot like us but gets much better results'. In Tucker, M.S. (Ed.) *Surpassing Shanghai: An agenda for American education built on the World's leading systems*. Cambridge, Mass.: Harvard Education Press. [Also published as part of: OECD (2011) 'Ontario, Canada: Reform to support high achievement in a diverse context'. In *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education*, OECD Publishing. Available online at: http://dx.doi.org/10.1787/9789264096660-en] (Accessed 16 December 2013).

Milburn, C. (2010) Wanted: a new respect for learning. *The Age*, 12 October 2010. Available online at: http://www.theage.com.au/national/education/

wanted-a-new-respect-for-learning-20100709-103do.html (Accessed 13 December 2013).

Nisbet, I. (2012) *What can our education system really learn from the East?* Presentation given in Cambridge, 20 April 2012. Cambridge: University of Cambridge International Examinations.

Oates, T (2010) *Could do better: Using international comparisons to refine the National Curriculum in England*. Cambridge: Cambridge Assessment. Available online at: http://www.cambridgeassessment.org.uk/Images/112281-could-do-better-using-international-comparisons-to-refine-the-national-curriculum-in-england.pdf (Accessed 16 December 2013).

Organisation for Economic Development (OECD) (2011a) 'Shanghai and Hong Kong: Two distinct examples of education reform in China'. In *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education*, OECD Publishing. Available online at: http://dx.doi.org/10.1787/9789264096660-en (Accessed 16 December 2013).

Organisation for Economic Development (OECD) (2011b) 'Lessons for the United States'. In *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education*. OECD Publishing. Available online at: http://dx.doi.org/10.1787/9789264096660-en (Accessed 16 December 2013).

Office for National Statistics (2013) 2011 *Census: QS103UK Age by single year, local authorities in the United Kingdom*. London: Office for National Statistics.

Pearson (2012) *Supporting Singapore's teachers*. Pearson, The Learning Curve. Available online at: http://thelearningcurve.pearson.com/videos?id=15 (Accessed 16 December 2013).

Pearson (2013a) *Talent for teaching: Singapore's National Institute of Education and its role in teacher development*. Pearson, The Learning Curve. Available online at: (http://thelearningcurve.pearson.com/case-studies/article/talent-for-teaching (Accessed 16 December 2013).

Pearson (2013b) *Raising the bar: In Australia, setting national standards for teachers is no easy task*. Pearson, The Learning Curve. Available online at: http://thelearningcurve.pearson.com/case-studies/article/raising-the-bar (Accessed 16 December 2013).

Peters, F. (2011) *The historical evolution of public education in Alberta*. Calgary, Alberta: The Sheldon Chumir Foundation for Ethics in Leadership. Available online at: http://www.youtube.com/watch?v=QpAYY9EGg0U (Accessed 16 December 2013).

Phillips, N. (2010) "Attitude" a Key Factor in China's Educational Success. *Zela Wela Kids: The financial life skills blog for families*, 10 December, 2010. Available online at: http://www.zelawelakids.com/blog/bid/49982/Attitude-a-Key-Factor-in-China-s-Educational-Success (Accessed 16 December 2013).

Prime Minister's Science, Engineering and Innovation Council (PMSEIC) (2009) *Transforming Learning and the Transmission of Knowledge: Preparing a learning society for the future*. Canberra: PMSEIC. Available online at: http://www.innovation.gov.au/Science/PMSEIC/Documents/TransformingAndTheTransmissionOfKnowledge.pdf (Accessed 16 December 2013).

Reynolds, D., & Farrell, S. (1996) *Worlds apart? A Review of International Surveys of Educational Achievement involving England*. London: Her Majesty's Stationery Office.

Schleicher, A. (2011) *Building a High-Quality Teaching Profession: Lessons from around the world*. International Summit on the Teaching Profession, 16-17 March 2011. Available online at: http://www.oecd.org/education/preschoolandschool/programmeforinternationalstudentassessmentpisa/47506177.pdf (Accessed 16 December 2013).

Slobodian, A. (2012) 'No zeros' case shows Alberta needs to set consistent grading policy, *Calgary Sun*, 11 September 2012. Available online at: http://www.calgarysun.com/2012/09/12/no-zeros-case-shows-alberta-needs-to-set-consistent-grading-policy (Accessed 16 December 2013).

Spar, D. (2009). What Higher Education can learn from the case of Singapore. *Forum for the future of higher education*. Available online at: http://net.educause.edu/ir/library/pdf/ff0905s.pdf (Accessed 16 December 2013).

Srilal, M. (1999) Culture-Singapore: 'Singlish' is Out in Re-education Campaign. *International Press Service*, 27th August 1999. Available online at: http://www.ipsnews.net/1999/08/culture-singapore-singlish-is-out-in-re-education-campaign/ (Accessed 16 December 2013).

Stewart, V. (2011) 'Singapore: A journey to the top, step by step'. In Tucker, M.S. (Ed.) *Surpassing Shanghai: An agenda for American education built on the World's leading systems*. Cambridge, Mass.: Harvard Education Press. [Also published as: OECD (2011) 'Singapore: Rapid improvement followed by strong performance'. In *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education*, OECD Publishing. Available online at: http://dx.doi.org/10.1787/9789264096660-en] (Accessed 16 December 2013).

Tan. C. (2012) The culture of education policy making: curriculum reform in Shanghai. *Critical Studies in Education*, **53**, 2, 153-167. Available online at: http://www.tandfonline.com/doi/abs/10.1080/17508487.2012.672333 (Accessed 16 December 2013).

Tucker, M.S. (2011) How the top performers got there: Analysis...and synthesis. In Tucker, M.S. (Ed.) *Surpassing Shanghai: An agenda for American education built on the World's leading systems*. Cambridge, Mass.: Harvard Education Press.

Watson, P. (2012) Singapore and its education reforms: but what about freedom? *Montrose 42's Blog*, 19 March 2012. Available online at: http://montrose42.wordpress.com/2012/03/19/singapore-and-its-education-reforms/ (Accessed 16 December 2013).

Wikipedia (2013a) *Education in Hong Kong*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/Education_in_Hong_Kong (Accessed 16 December 2013).

Wikipedia (2013b) *Cram schools in Hong Kong*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/Cram_schools_in_Hong_Kong (Accessed 16 December 2013).

Wikipedia (2013c) *Working time*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/Working_time (Accessed 16 December 2013).

Wikipedia (2013d) *Singapore in Malaysia*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/Singapore_in_Malaysia (Accessed 16 December 2013).

Wikipedia (2013e) *Education in Victoria*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/Education_in_Victoria (Accessed 16 December 2013).

Wikipedia (2013f) *History of education in New Zealand*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/History_of_education_in_New_Zealand (Accessed 16 December 2013).

Wikipedia (2013g) *Think Big*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/Think_Big (Accessed 16 December 2013).

Wikipedia (2013h) *National Certificate of Educational Achievement*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/National_Certificate_of_Educational_Achievement (Accessed 16 December 2013).

Wikipedia (2013i) *Culture of New Zealand*. Wikipedia, The Free Encyclopedia. Available online at: http://en.wikipedia.org/wiki/Culture_of_New_Zealand (Accessed 16 December 2013).

Yam, J. (2010) *Singapore Education System – Comparison, Issues and Challenges*. Dr John Yam Blog. Available online at: http://drjohnyam.blogspot.co.uk/2010/11/singapore-education-system-comparison.html (Accessed 16 December 2013).

Yeung, A.S. & Yeung, A. (2008) 'Ability versus effort: perceptions of students from the East and from the West.' In Tan, O.S., McInerney, D.M., Liem, A.D. & Tan, A. (Eds.) *What the West can learn from the East: Asian perspectives on the psychology of learning and motivation*. Charlotte, North Carolina: Information Age.

Examining the impact of tiered examinations on the aspirations of young people

Tom Benton Research Division

Introduction

Tiered examinations are commonly employed within the General Certificate of Secondary Education (GCSE) examinations in the UK. Within a given subject, tests at different levels of difficulty are developed and then teachers or schools can decide which tier is most appropriate for their candidates (Dhawan and Wilson, 2013). Within current tiered GCSEs, more able candidates will be allocated to more difficult, "higher" tier tests whereas less able candidates will be directed towards the "foundation" tier. The highest GCSE grades (A*–B) are only available to those candidates who take the higher tier version of the test. In the past, GCSE Maths used a three tier structure where pupils of low, medium and high ability were directed towards foundation, intermediate and higher tier versions of examinations.

The aim of tiering is to ensure that the difficulties of exam papers are correctly tailored to the ability of the candidates taking them; this should ensure more accurate measurements and also a better experience for candidates as they do not spend time addressing questions that are either too easy or too difficult given their level of skill. However, tiered examinations have been criticised for potentially damaging the aspirations of young people. For example, the Department for Education's (DfE) 2012 consultation into the reform of qualifications stated:

> *The prospects for those students taking a foundation tier paper are poor… Having a grade-cap in foundation tier examinations is also likely to be de-motivating and limit the aspirations of students*.
> (DfE, 2012).

Other research has linked the use of tiered examinations with the more general issue of ability setting and, similarly, suggested that, in this context, tiered examinations may have a demotivating effect (Boaler, 1997; Boaler *et al*. 2000). For example, Boaler (1997) reported results from qualitative research in one school. She found that:

> *… students became disillusioned and demotivated by the limits placed upon their achievement within their sets*. (Boaler, 1997).

In the light of these statements, the aim of this paper is to provide a large scale, quantitative examination of the extent of the link between GCSE entry tier and aspirations and also to investigate the extent to which this link can be explained by differences in achievement and background characteristics of pupils. It should be noted that, in some sense, there is an intrinsic link between aspirations and entry tier in that, in general, students can only continue to higher level study within a subject if they achieve a grade B or above, and this can only be achieved if they enter the higher tier. Thus, by entering the lower tier, the decision not to continue studying the given subject further beyond GCSE has already been made. As such, it is not sensible to quantitatively examine the link between tiers and aspirations *within* a given subject. However, the quotes above hint at a wider form of de-motivation and disillusionment coming from students being placed in a lower tier, suggesting that being entered for such an examination may harm students' educational aspirations and desire for learning across all subjects, not just the subject they are entered for. It is this hypothesis that is explored in this paper. Namely, we examine whether there is any evidence of entering candidates for lower tier examinations having a negative impact on their wider educational aspirations or, indeed, on their chances of participating in post-compulsory education.

It should be noted that this paper does not explore the effects of tiered assessment on the achievement of young people but is purely concerned with the effect on aspirations. Furthermore, this paper only examines the possible effects of tiering during Key Stage 4. Any effects of tiering on pupils prior to the beginning of Key Stage 4[1] are beyond the scope of this research.

Data and Method

The research makes use of data available from the Longitudinal Study of Young People in England[2] (LSYPE). The LSYPE began collecting data on the attitudes of around 16,000 Year 9 pupils in a representative sample of English schools in 2004. These pupils have been followed up in every subsequent year so that data has been collected on their educational and attitudinal development over time. Of particular focus for this paper is data regarding the entry tier of these young people in their GCSEs; the majority of which were taken in summer 2006 and are recorded in the National Pupil Database (NPD). By linking this data to questionnaire responses about young people's future educational aspirations, we can explore the relationships between GCSE entry tiers and aspirations.

Data on the entries and achievements at Key Stage 4 of the young people participating in the LSYPE is available from the NPD. For every qualification taken by young people during Key Stage 4, a number of details including qualification type, subject and achieved grade are recorded. Also recorded is a qualification identifier provided by the exam board delivering the qualification. For GCSEs delivered by AQA (and occasionally OCR[3]) the qualification identifier is suffixed by the letters "F", "I" or "H" to indicate whether the candidate took the qualification at the foundation, intermediate or higher tier respectively. Using this information, for a sub-sample of young people, it was possible to identify the tier at which they were entered for their Maths and English GCSEs.

Data on the educational aspirations of these young people is available from a questionnaire completed during 2006 by around 12,000 of the

1. At the time of the data collection, pupils were also entered for higher and lower tiers in Key Stage 3 tests.
2. For further detail on this study please visit: https://www.education.gov.uk/ilsype/workspaces/public/wiki/Welcome/LSYPE
3. But never Edexcel or WJEC.

LSYPE participants at the end of Year 11. This questionnaire asked specific questions about whether young people intended to stay in education post-16 and also about how likely it was that they would apply to university in the future. The aim of analysis was to explore the relationship between young people's responses to these questions and their GCSE entry tiers in Maths and English.

In undertaking such an analysis it is immediately clear that any link between tier and educational aspirations could be explained by a number of pre-existing external factors. For example, it may be that pupils who enter lower tiers at GCSE are those that had low aspirations to begin with and so would be likely to continue to be those with low aspirations even if the GCSE tier itself had no negative effect. Alternatively, it is extremely likely that pupils entering lower tiers at GCSE will be those with lower levels of ability on average and thus would tend to have lower aspirations regardless of their entry tier. For these reasons it was important for the analysis to take account of these factors and others in order to make valid conclusions about the relationship between entry tier and aspirations. The analysis accounted for the following potentially influential factors:

- Educational ability
- Gender
- Eligibility for free school meals
- Level of special educational needs
- Ethnicity
- Language spoken at home
- Initial intentions regarding post-16 education as measured in Year 9[4]
- Feelings in Year 9 about likelihood of applying for, and being accepted into, university in the future
- Attitude to school work as measured in Year 9 using a composite score derived from 12 survey questions[5]
- Number of risk factors experienced by students[6] in Year 9

Analysis comparing the aspirations of pupils in each tier was undertaken using a combination of propensity score matching and multilevel modelling. Initially pupils were divided into two groups based upon their entry tier. Pupils whose entry tier was not identified were removed from analysis. Within each group, pupils with background characteristics unlikely to be found in the opposing group were removed from analysis. For example, because very high attaining students were unlikely to be entered for lower tier exams, all such pupils were removed from the data set. At this point an initial comparison between the aspirations of the young people in each tier was made. Responses from the group of students in the higher of the two tiers being compared were weighted according to the background characteristic of students. This was done such that, after weights were applied, the background characteristics of pupils in the higher tier matched the background characteristics of those in the lower tier. Comparing aspirations between lower tier pupils, and the resulting weighted data for higher tier pupils, provided an estimate of the differences between the two groups whilst accounting for the effect of other influential factors. The statistical significance of differences was then verified using multilevel modelling.

4. That is, prior to beginning study for GCSEs.

5. See page 381 of http://www.esds.ac.uk/doc/5545/mrdoc/pdf/5545wave_one_documentation.pdf for further details.

6. Risk factors include involvement in activities such as smoking, alcohol or drug abuse, vandalism, truancy and others.

For the purposes of analysis, educational ability was measured in each of two ways; either using Key Stage 3 attainment[7] or Key Stage 4 attainment[8]. In the latter case, because GCSE entry tiers restrict the grades available to students, this placed a restriction on the data that could be meaningfully included in analysis. For English GCSE, aspirations of foundation and higher tier pupils could only be meaningfully compared for those achieving grade C or D in GCSE English. For Maths GCSE, foundation and intermediate tier pupils could only be meaningfully compared for those achieving grade D or E, whereas intermediate and higher tier pupils could only be compared for those achieving grade B or C. No such explicit restrictions were placed on the analyses which used Key Stage 3 attainment to account for differences in the educational ability of students within different tiers[9].

As noted earlier, entry tier was only identifiable for candidates taking their GCSEs with particular exam boards. For English GCSE, because AQA is the major provider of this qualification, all relevant data could be identified for a sample of over 7,000 pupils. However, for Maths GCSE, because a greater proportion of candidates take the subject with Edexcel, a sample of less than 3,000 pupils was available for analysis. Furthermore the data for Maths GCSE was split across three tiers rather than two. For this reason estimates of the relationship between Maths entry tier and aspirations are subject to greater uncertainty than similar estimates based on entry tier in English[10].

Results

Results of analysis comparing pupils entered for different tiers whilst controlling for attainment at Key Stage 3 and other background factors are shown in Table 1. The first two columns of data show, for the young people in each tier retained within the analysis, the percentage saying that they intend to stay in education post-16 and the percentage saying they are likely to apply to university in future. The third column then shows the adjusted figure for higher tier candidates after weighting the data to account for the background characteristics of these young people. For example, the first row of data shows that 82 per cent of candidates entering foundation tier English intended to stay in education post-16 compared to 95 per cent of higher tier candidates. However, weighting the data to account for background characteristics reduces the figure for higher tier candidates to 87 per cent. In other words, this means that we estimate that a group of candidates with background characteristics equivalent to those who entered the lower tier, but who actually entered the higher tier would have an 87 per cent chance of saying they intend to stay in education post-16. The final two columns of data present the number of pupils available for analysis within each comparison. A graphical presentation of the same analysis is shown in Figure 1.

These results show that although there is a strong relationship between GCSE entry tier and educational aspirations, much of this link is

7. As measured by fine graded point scores in each subject.

8. Measured by the grade achieved in the subject of interest as well as the "capped total points score" which provided a more general measure of pupils' attainment across all their Key Stage 4 subjects.

9. Although, due to the very strong association between Key Stage 3 attainment and entry tier, a number of pupils with achievement levels that were not comparable across tiers were removed from analysis.

10. Another impact of the smaller sample size for analysis of Maths GCSE was that, for analysis taking account of Key Stage 3 attainment, it was not possible to adequately match higher and lower tier candidates across all of the listed background characteristics. For this reason it was necessary to restrict analysis to take account of only: Key Stage 3 attainment, gender, prior intentions regarding post-16 education and prior attitudes to university.

Table 1: Differences in aspirations between candidates entering different tiers before and after accounting for differences in Key Stage 3 attainment and other background characteristics

GCSE subject (tiers being compared)	*Outcome*	*Lower tier*	*Higher tier*	*Higher tier (weighted)*	*N (Lower tier*	*N (Higher tier)*
English (foundation vs higher)	% Intending to stay in education post-16	82	95	87	*2920*	*4222*
	% Likely to apply to HE	47	80	53	*3041*	*4273*
Maths (foundation vs intermediate)	% Intending to stay in education post-16	79	89	80	*558*	*1553*
	% Likely to apply to HE	36	65	55	*582*	*1591*
Maths (intermediate vs higher)	% Intending to stay in education post-16	89	97	88	*1255*	*607*
	% Likely to apply to HE	68	88	77	1286	613

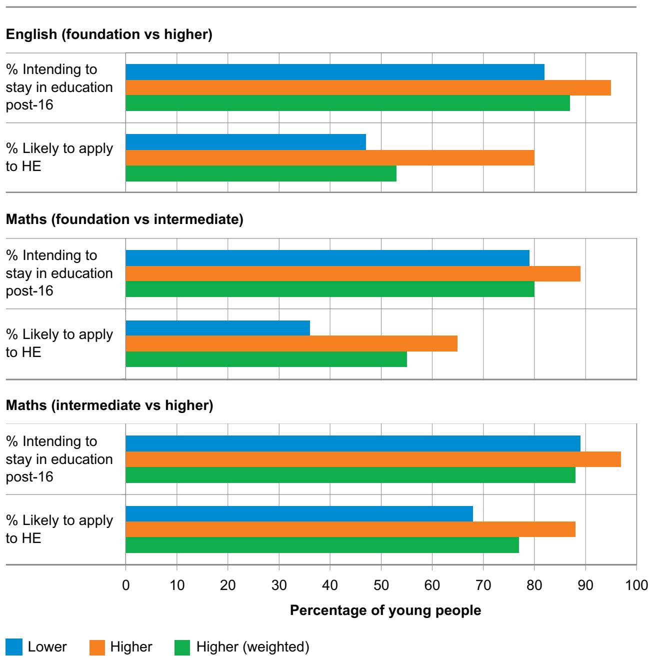


Figure 1: Differences in aspirations between candidates entering different tiers before and after accounting for differences in Key Stage 3 attainment and other background characteristics

explained by the background characteristics of young people. In particular, once the impact of background characteristics has been accounted for, there appears to be very little difference between young people entered for lower and higher tiers in terms of their intentions to remain in education post-16. Having said this, the difference between tiers for English GCSE remains statistically significant, albeit small.

11. Despite the apparently large size of the difference in aspirations regarding higher education between intermediate and higher tier Maths students, the relatively small sample size available for this analysis means that this difference is not found to be statistically significant.

Table 2: Differences in aspirations between candidates entering different tiers before and after accounting for differences in Key Stage 4 attainment and other background characteristics

GCSE subject (tiers being compared)	*Outcome*	*Lower tier*	*Higher tier*	*Higher tier (weighted)*	*N (Lower tier*	*N (Higher tier)*
English (foundation vs higher)	% Intending to stay in education post-16	86	92	88	*2090*	*1722*
	% Likely to apply to HE	54	68	55	*2156*	*1754*
Maths (foundation vs intermediate)	% Intending to stay in education post-16	79	84	81	*402*	*541*
	% Likely to apply to HE	38	52	41	*416*	*569*
Maths (intermediate vs higher)	% Intending to stay in education post-16	93	96	93	*869*	*442*
	% Likely to apply to HE	76	87	76	*879*	*447*

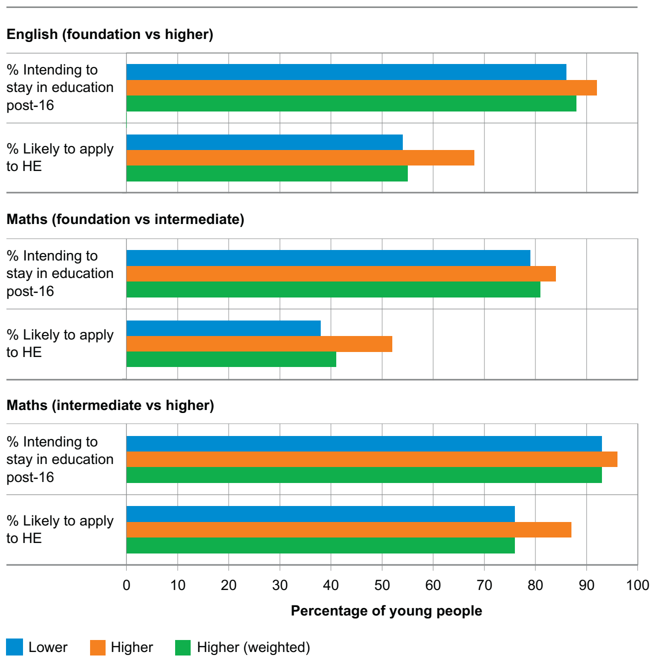


Figure 2: Differences in aspirations between candidates entering different tiers before and after accounting for differences in Key Stage 4 attainment and other background characteristics

Furthermore, there are statistically significant differences between tiers in terms of aspirations regarding Higher Education (HE) for both English GCSE and for Maths GCSE when comparing those in the foundation tier to those in the intermediate tier[11].

A possible criticism of the above analysis is that it does not adequately take account of the main factor likely to determine the entry tier of young people; namely their ability in the given subject at the time at which they were entered for the exam. To address this, the same analysis was repeated but taking account of achievement at Key Stage 4 rather than

Table 3: Differences in probability of continuing in education between candidates entering different tiers before and after accounting for differences in Key Stage 3 attainment and other background characteristics

GCSE subject (tiers being compared)	*Outcome (% within education or training in....)*	*Lower tier*	*Higher tier*	*Higher tier (weighted)*	*N (Lower tier*	*N (Higher tier)*
English (foundation vs higher)	October 2006	79	94	81	*2748*	*4015*
	October 2007	70	88	71	*2502*	*3807*
Maths (foundation vs intermediate)	October 2006	76	88	84	*503*	*1474*
	October 2007	67	79	73	*468*	*1369*
Maths (intermediate vs higher)	October 2006	88	97	93	*1215*	*589*
	October 2007	81	92	89	*1118*	*573*

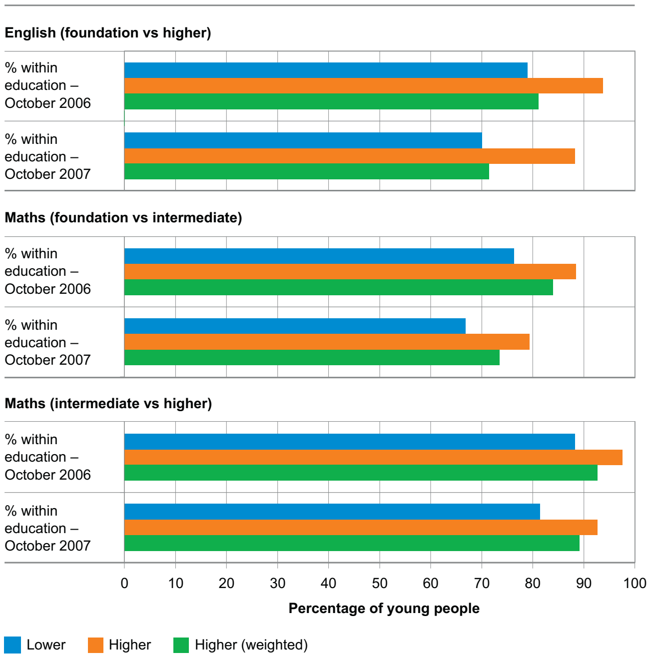


Figure 3: Differences in probability of continuing in education between candidates entering different tiers before and after accounting for differences in Key Stage 3 attainment and other background characteristics

Key Stage 3. The results of this analysis are shown in Table 2 and Figure 2.

As with the previous analysis, these tables show that, before taking account of the impact of background characteristics, there are some large differences in the educational aspirations of young people. However, once the abilities and characteristics of the different students are taken into account, these differences in aspirations almost entirely vanish, Indeed, none of the differences between tiers shown in Figure 2 are statistically significant once we have taken the impact of other factors into account. This implies that, all else being equal, it does not matter whether a candidate achieves a grade C (for example) in the higher tier or the lower tier; the future aspirations of the student will be identical. This would imply that students should be entered for the tier most appropriate to their ability, and there is no need for concern that such a strategy may damage their educational aspirations.

A potential criticism of this approach is that it could be argued that entry tier affects aspirations by first reducing the likely achievement of young people at GCSE. Thus, controlling for attainment within GCSE itself

Table 4: Differences in probability of continuing in education between candidates entering different tiers before and after accounting for differences in Key Stage 4 attainment and other background characteristics

GCSE subject (tiers being compared)	*Outcome (% within education or training in....)*	*Lower tier*	*Higher tier*	*Higher tier (weighted)*	*N (Lower tier*	*N (Higher tier)*
English (foundation vs higher)	October 2006	83	87	82	*1952*	*1572*
	October 2007	73	79	72	*1809*	*1450*
Maths (foundation vs intermediate)	October 2006	79	82	78	*369*	*530*
	October 2007	67	71	65	*349*	*482*
Maths (intermediate vs higher)	October 2006	91	97	96	*835*	*429*
	October 2007	85	90	82	*781*	*414*

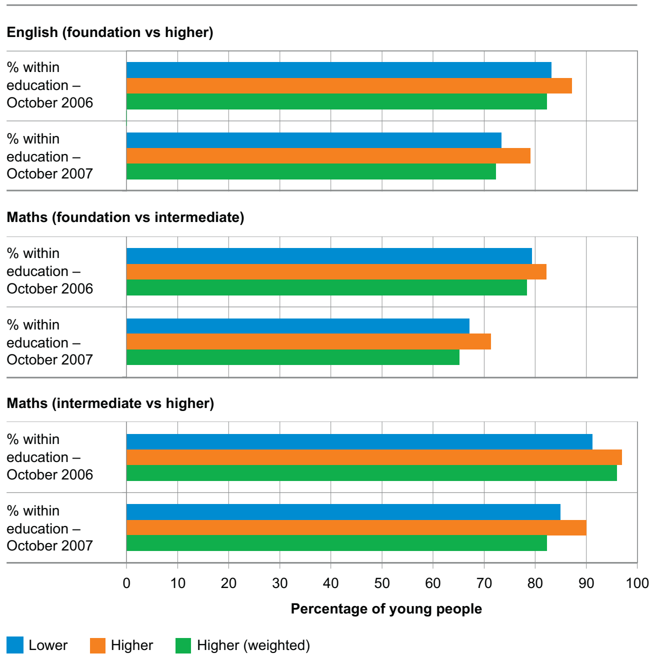


Figure 4: Differences in probability of continuing in education between candidates entering different tiers before and after accounting for differences in Key Stage 4 attainment and other background characteristics

is inappropriate. However, our earlier analysis has shown that even if we only control for attainment at Key Stage 3, much of the difference between the aspirations of candidates in different tiers can be explained. For this reason we can conclude that the impact of GCSE entry tier on educational aspirations is quite small at worst and, when we allow for the possible impact of other potential explanatory variables not included within this analysis, potentially non-existent.

Although examining the association between tiers and aspirations is of some value in its own right, aspirations do not necessarily translate into actual continuation in education (Gorard *et al*., 2012). That is, just because a pupil intends to do something doesn't necessarily mean that they actually will. For this reason, it was of interest to also examine the relationship between tiers and the actual educational destinations of pupils at the start of each of the academic years after the end of compulsory education. That is, whether they were participating in education (including apprenticeships) in October 2006 and October 2007.

The same analysis as for aspirations was undertaken this time with the outcome of interest being whether young people were participating in any form of education in October 2006 and October 2007. The results after taking account of background variables including Key Stage 3 attainment are shown in Table 3 and Figure 3. The results after taking account of Key Stage 4 attainment are shown in Table 4 and Figure 4. The findings with respect to actual destinations are in line with those described earlier with respect to aspirations. Before taking account of the background characteristics of young people there is a clear difference in the probability of those entered for different tiers remaining in education post-16. However, once the influence of background characteristics is taken into account this difference is greatly reduced. Furthermore, as shown in Table 4 and Figure 4, once we account for the achievement of pupils at Key Stage 4 there is essentially no difference between the educational destinations of those who were entered for the lower tier and those entered for the higher tier.

Summary and caveats

The analysis presented here has explored the link between entry tier in Maths and English GCSE and future educational aspirations as measured within the Longitudinal Study of Young People in England (LSYPE). The analysis shows that any differences in aspirations or, indeed, chances of actually continuing in post-compulsory education can be entirely explained by the background characteristics of young people and in particular their educational ability as measured by their level of achievement at Key Stage 4. Whilst it could be argued that taking account of achievement at Key Stage 4 is inappropriate (as this could itself be affected by entry tier), our analysis has also shown that even taking account of achievement at Key Stage 3 is sufficient to explain much of the difference between higher and lower tier students.

It should be noted that this analysis is based on somewhat old data; the young people being studied completed their GCSEs in 2006. Furthermore, because information about entry tier is only available from particular exam boards, analysis is largely restricted to pupils taking Maths and English with AQA rather than with any other exam boards. Thus our analysis implicitly assumes that the impact of tiering will be similar across different exam boards.

Nevertheless, despite the need to restrict to candidates entering English and Maths to particular exam boards, we have successfully been able to compare the educational aspirations of several thousand higher and lower tier candidates. Once differences in the characteristics of these pupils are accounted for, we have seen remarkable similarity in their educational aspirations. This provides a clear empirical challenge to the statement that being placed in a lower tier examination will lead to demotivation and disillusionment. How teachers and schools should decide upon the most appropriate tier for their candidates remains an open question. However, it is clear that this decision can be made without fear that entering students for a lower tier will have wide reaching consequences beyond the individual GCSE subject.

References

Boaler, J. (1997). Setting, social class and survival of the quickest. *British Educational Research Journal*, **23**, 575–595.

Boaler, J., William, D., & Brown, M. (2000). Students' experiences of ability grouping – disaffection, polarisation and the construction of failure. *British Educational Research Journal*, **26**, 631–648.

Dhawan,V. & Wilson F. (2013). Comparing difficulty of GCSE tiered examinations using common questions. *Research Matters: A Cambridge Assessment Publication*, **16**, 49–56.

Department for Education (DfE) (2012) *Reforming Key Stage 4 Qualifications*. Consultation document issued on 17th September 2012.

Gorard, S., See, B. H. & Davies, P. (2012). *The impact of attitudes and aspirations on educational attainment and participation*. Joseph Rowntree Foundation: York.

Education and neuroscience

Vikas Dhawan Research Division

If we value the pursuit of knowledge, we must be free to follow wherever that search may lead us. Adlai E. Stevenson Jr. (1952)

Introduction

This study was aimed at exploring how recent developments in neuroscience (the study of the structure and functioning of the brain) might affect the fields of education and test development in the future.

The study investigated some of the potential areas of application as well as limitations of neuroscience in education. A brief summary of the application of neuroscience in some other areas is also given. These are marketing and advertising, health, psychology and politics.

The main findings of this study were:

- There is a growing interest in the media, commercial organisations and the education sector for anything related to neuroscience.
- Various universities and academic institutions have started centres for research in neuroscience and education including Cambridge, Oxford, Bristol, University College London (UCL), Birkbeck, Harvard and Stanford.
- The field of health and medicine is leading the research in neuroscience which is being used in other fields.
- Neuroscience applications are in great demand in consumer marketing and advertising.
- Considerable research is being carried out in understanding learning disabilities (such as dyslexia and dyscalculia) using neuroscience.

- New classroom-teaching approaches based on neuroscience are becoming increasingly popular.
- Caution needs to be observed on claims made for applications of neuroscience – not all activity is scientifically valid and there are many 'neuromyths' floating around (e.g. we use only 10% of our brain).
- The techniques used in mapping the activity of brain (such as scanning) are expensive and cumbersome at this stage and therefore not suitable for large scale testing.
- The applicability of neuroscience for developing or validating educational assessments at present appears limited.
- More neuroscience-based applications are likely to emerge in the near future especially as various governments are committing to research in this field – the USA has announced investment of $3 billion over 10 years in neuroscience research and the European Commission has recently awarded €1 billion to the Human Brain Project under its Future and Emerging Technologies initiative.
- Educational authorities and awarding organisations should keep themselves abreast of how neuroscience might lead to innovative teaching and test-development practices.

What is neuroscience?

Neuroscience implies the study of the working of the brain. This field has been growing significantly in recent years. Understanding of how the human brain works is being increasingly applied to various fields such as health, psychology, education, marketing, politics and law. In all such applications the objective is to provide solutions based on the underlying causes of why and how human beings function the way they do. Deciphering the brain would indeed be the 'holy grail' in designing solutions in all walks of life. It could minimise the dependence on individual biases such as judgement and perception by making available a more reliable source of information instead – the functioning of the brain. For instance, if we have a precise understanding of how learning difficulties manifest in the brain there is a better chance of providing more targeted solutions. Similarly, we might be able to get a better picture of the psychological state of an individual by enhancing our understanding of how different behaviours are represented in the brain. Neuroscience might also allow us to develop educational tests that tap the skills that we intend to assess in a more targeted and effective manner.

Neuroscientists use the term 'mapping' to study the structure and functioning of the brain. A major project to map the brain was launched by the National Institutes of Health in the USA in 2009 (Connectome, 2013). The project aims to prepare a network map of the brain by using images of the brain and relate it with behavioural tests. Data and results from this project have now been made freely available to the scientific community for further research. USA President Barack Obama has recently announced plans to invest $3 billion in neuroscience research over a period of ten years (TES, 2013; The New York Times, 2013; BRAIN Initiative, 2013). The project is known as the BRAIN Initiative (Brain Research through Advancing Innovative Neurotechnologies), also referred to as the Brain Activity Map project. It is aimed at building a comprehensive activity map of the brain – similar to what the Human Genome Project (1990–2003) did for genetics and came with a price tag of $3.8 billion. The European Commission has recently awarded €1 billion to the Human Brain Project under its Future and Emerging Technologies initiative. (Human Brain Project, 2013; Europa, 2013).

Over the next decade the project will aim to develop a large scale information and communications technology (ICT) infrastructure specifically for understanding the brain.

Brain mapping techniques

The major techniques used in brain mapping (Mapping Techniques, 2013a; 2013b) are:

- Computer axial tomography (CAT) scan: X-rays of structures of the brain from different angles.
- Electroencephalography (EEG): measures brain activity using detectors implanted in the brain or worn on a cap.
- Functional Magnetic Resonance Imaging (fMRI): shows images of brain activity while subjects work on various tasks.
- Magnetoencephalography (MEG): records brain activity by using electrical currents occurring naturally in the brain.
- Positron emission tomography (PET) scan: produces three-dimensional images of radioactive markers in the brain.

The techniques are meant to capture changes in neural activity due to specific demands placed on the brain by various tasks. The participants might either be healthy individuals or those suffering from a disorder, depending on the study. Out of the techniques mentioned above, fMRI and EEG are being extensively used by researchers to study brain activity.

Figure 1 shows an MRI scanning machine. The participant lies inside the scanner and is given audio or visual input.

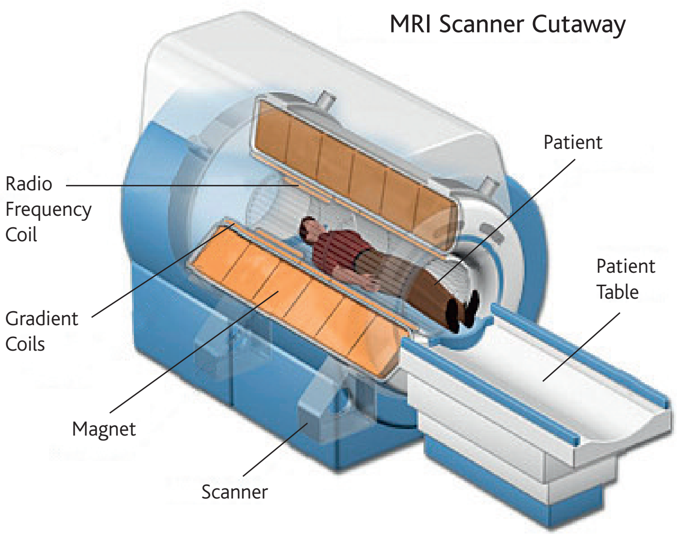


Figure 1: Scanner used for fMRI studies (Magnet Lab, 2013).
Image courtesy of Magnet Lab – National High Magnetic Laboratory, Florida State University.

Figure 2 gives an example of fMRI output. The images have been taken from a study which compared children aged 11 on a task of writing a letter (Todd *et al*., 2011). The figure shows the difference in the activity patterns in the brains of children who were classified as good writers and poor writers.

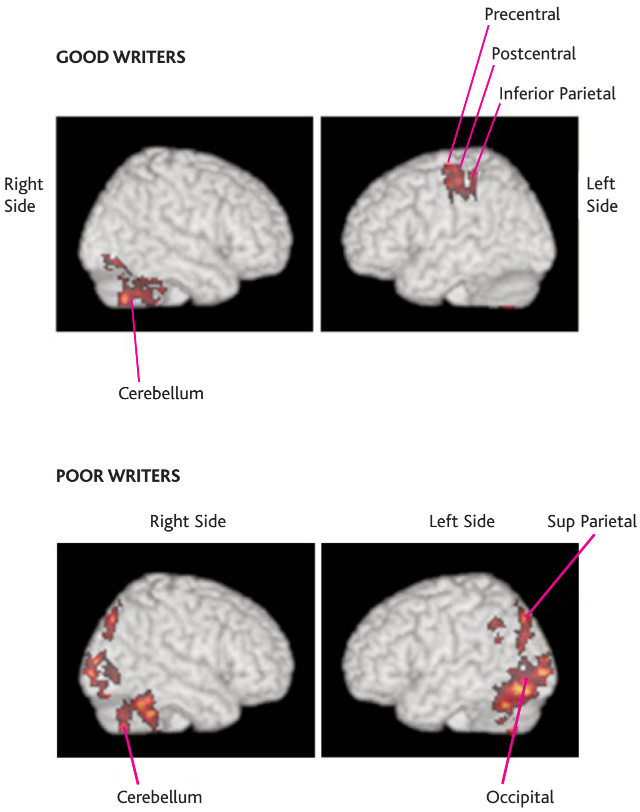


Figure 2: Example of fMRI scan images of brain: comparison of brain activity in children classified as good and poor writers (Todd *et al*., 2011).
Image courtesy of Springer©, Part of Springer Science+Business Media

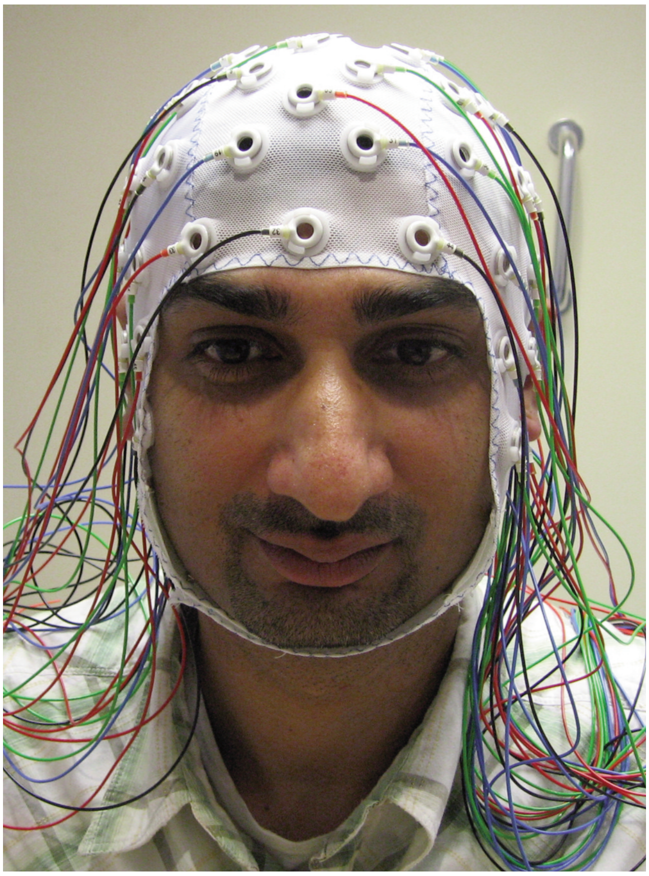


Figure 3: Example of an EEG recording.
Image courtesy of the National Institute for Health (NIH) Research: Nottingham Hearing Biomedical Research Unit.

Figure 3 gives an example of the EEG technique. The participant is required to wear a cap on their skull which is then fitted with electrodes, which can take up to an hour. The electrical activity of the brain is recorded according to the study requirements – while engaged in some activity or in a resting state.

Applications of neuroscience

There has been huge interest recently in the application of neuroscience to various fields. The flurry of activity and media interest has given rise to many new terms such as neuromarketing, neuroeconomics, neuroaesthetics, neurotheology, neurolaw, neueroanthropology and neuropolitics (Legrenzi, Umilta and Anderson, 2011).

Marketing

One of the most commercially intensive applications of neuroscience is neuromarketing (also called consumer neuroscience). It involves studying which products/packaging/ideas 'appeal' to the brain and devising marketing strategies accordingly. Techniques such as fMRI and EEG are used along with biometric measures such as eye movements, heartbeat and skin response, coupled with interviews/questionnaires to understand the subconscious preferences of consumers. A large number of market research organisations, including the well-known ones such as Ipsos, Gallup and Nielsen, now offer a neuroscience component in their portfolio of research tools. Major companies such as Google, HP, Microsoft and Coca-Cola are known to use neuroscience and biometric measures for devising their marketing strategies. According to one UK-based estimate, more than 10 per cent of prime time TV advertisements have been developed using neuromarketing techniques (Guardian, 2012).

Some of the findings claimed by NeuroFocus, probably one of the world's leading neuromarketing firms (now a part of Neilsen), are:

- People prefer items with rounded edges than those with sharp corners;
- Mannequins and photos with missing heads turn consumers off;
- Men typically respond to a product's features whereas women are more interested in getting a deal.

(ABC News, 2011).

They also claim to have noticed some other gender differences (Neurorelay, 2012). For example, for insurance products, women reacted much more strongly than men to the character of the spokesperson, while men reacted to the price. In light-hearted adverts for snack foods, men reacted to slapstick humour, while women ignored it. In automotive adverts, men were interested only in the performance of the vehicle whereas women were interested in storage capacity and safety features. In spite of the widespread use of neuromarketing, the actual applications are not frequently published and it appears that companies prefer to keep them shrouded. Some of the applications designed by NeuroFocus for its clients are given in Figure 4.

Interestingly, film producers also seem to have started using neuroscience in film-making and marketing. For instance, it is reported that the trailers for the popular film, *Avatar*, were designed using viewer responses to different trailer scenes and sequences (Neurocinema, 2011).

In all these applications it is not clear if similar results could have been obtained by using only questionnaires and interviews without the need for the neuroscience component. The neuroscience experiments are

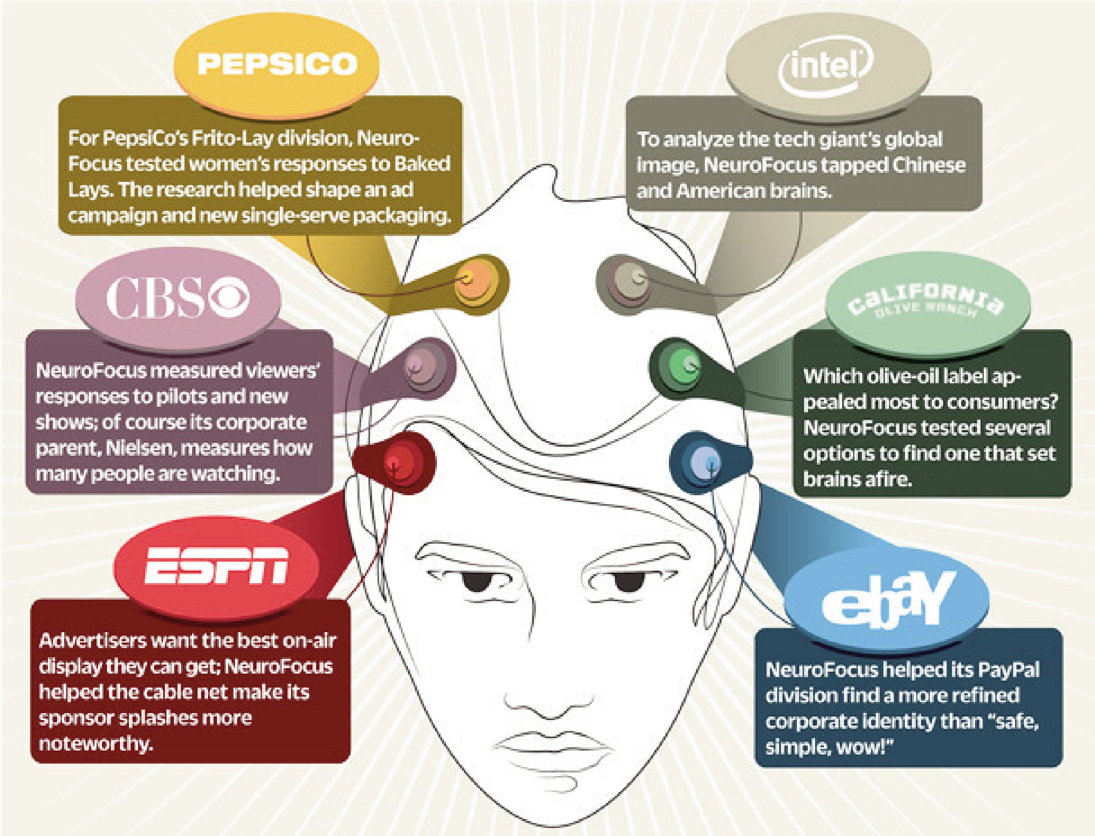


Figure 4: Examples of neuromarketing by NeuroFocus (Fast Company, 2011).
Image courtesy of SUPEREXPRESSO.

normally very expensive and their use would be beneficial only if they can provide additional evidence compared to the traditional methods such as interviews and focus groups.

Politics

Another area where neuroscience is being used is politics and law. It is well known that the 2012 Obama presidential campaign was heavily data-driven and used insights from behavioural economics and neuroscience to influence voters and to improve voter turnout (New York Times, 2012). Recent books such as The Victory Lab: The Secret Science of Winning Campaigns (Issenberg, 2012) and The political brain: The role of emotion in deciding the fate of the nation (Westen, 2007) also highlight this trend of how knowledge about human decision-making and the increasing power of analytics is being applied to political marketing.

The following USA-based study gives an example of using neuroscience research in this field. The study investigated if political awareness could be distinctly represented in the brain. Differences in brain activity were found between college students who were politically knowledgeable about Democrats and Republicans against those who did not know much about national politics (Schreiber, 2007 cited in Fowler and Schreiber, 2008). Figure 5 (Part A) shows that there were differences in the fMRI scans of those who were politically aware against those classified as political novices when asked questions about national politics. Part B in the figure shows the level of activation of specific brain regions for the two groups.

Health

The major developments in neuroscience emanate from the field of health and medicine. The Medical Research Council (MRC) Cognition and Brain Sciences Unit (CBU) in Cambridge conducts research in fundamental human cognitive processes such as attention, language, memory, and emotion using a combination of behavioural experiments, neuroimaging and computer modelling. The CBU works in close collaboration with the University of Cambridge and the local Addenbrooke's hospital.

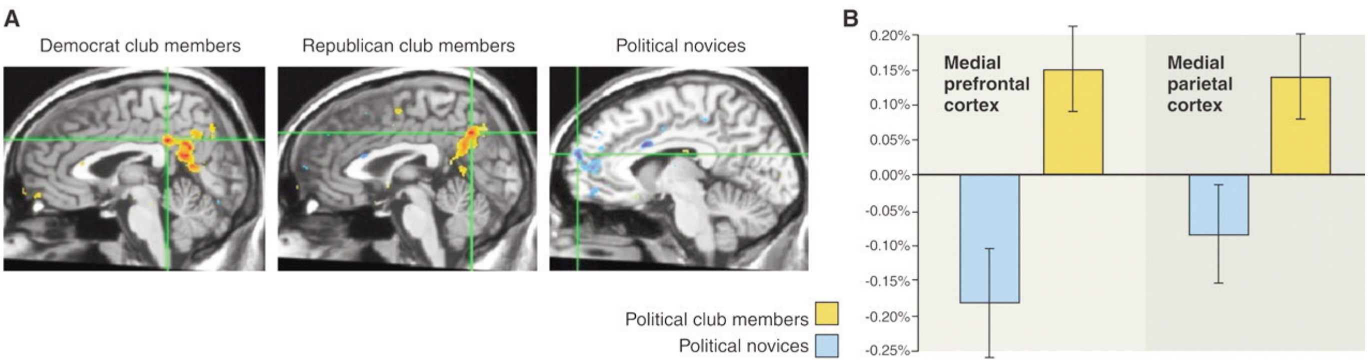


Figure 5: Difference in brain activity between politically knowledgeable against politically novice college students in the USA (Fowler and Schreiber, 2008).
Image courtesy of the authors.

The Center for Interdisciplinary Brain Sciences Research (CIBSR) at the Stanford University School of Medicine is a good example of multidisciplinary research. Its website states that the centre "is dedicated to research that will improve the lives and well-being of individuals with disorders of the brain and improve knowledge of healthy brain and behavioral development." CIBSR (2013).

The Center, which brings together experts from the fields of psychiatry, neurology, psychology, computer science, biostatistics and genetics, has developed a battery of assessment tools for measuring neurological and behavioural functions. The suite of assessments called the NIH[1] Toolbox measures motor, cognitive, sensory and emotional functions. It is available online, royalty-free and can be used by researchers and clinicians. The idea was to develop a 'common currency' or 'gold standards' against which individual performances can be compared across different neurological research studies. Education and psychology researchers might find this resource useful for their work in various domains such as cognitive psychology, emotional intelligence, marking and judgement processes and, also, impact of new modes of assessment such as computer-based tests – on both participants as well as judges. The NIH Toolbox could help provide a more targeted insight into understanding the level of stress and cognitive workload on markers in various modes of assessment, paper-based against computer-based.

Psychology

The assessment of personality traits could become more robust if a neurological basis of the traits can be established. Currently, most personality assessments are self-report measures, (that is, the participants answer statements about themselves and the responses are then used to estimate their personality profiles). Such self-report measures are susceptible to individual bias and social impression management. Using neuroscience techniques might help us gain a more uncluttered insight into personality and behaviour. DeYoung *et al*. (2010) found that four of the Big-Five[2] personality traits varied with the volume of different brain regions. The participants (n=116) were administered the self-report version of the Revised NEO Personality inventory (NEO-PI-R) (Costa and McCrae, 1992) which is based on the Big-Five model, followed by MRI scans. It was reported that E*xtraversion* co-varied with the brain region involved in processing reward information, N*euroticism* with regions associated with threat, punishment and negative affect, A*greeableness* with regions that process information about the intentions and mental states of other individuals and C*onscientiousness* with the region involved in planning and voluntary control of behaviour. No significant evidence was found for O*penness*. Figure 6 shows the association of different areas of brain with the personality traits. Research such as this could be used to develop more effective personality assessments.

An example of the application of neuroscience to personality assessment is the PRISM Brain Mapping© tool (Prism, 2013). It is an online personality assessment claimed to be based on neuroscience and can be used to identify the behavioural preferences that directly relate to personal relationships and work performance. The role of the instrument, as advertised, is "to explain behaviour in terms of the activities of the brain – how it marshals its billions of individual nerve cells to produce behaviour, and how those cells are influenced by the environment". We need to note here that questionnaires might not be different regardless of the technique used – traditional personality assessment or neuroscience. The insights gained into the personality of individuals using these two methodologies might also be similar to each other.

Education and neuroscience

A new area of research that has been gaining an increasing amount of interest is educational neuroscience which, as the name suggests, involves using neuroscience techniques in the field of education and learning. Another popular name which is being used to denote this field is Mind, Brain and Education (MBE). The field of educational neuroscience is vast and multi-disciplinary with perhaps no clear definition as yet. Szücs and Goswami (2007) define it as: "the combination of cognitive neuroscience and behavioral methods to investigate the development of mental representations"[4].

Most of the studies in this area involve using brain-scanning techniques, including fMRI and EEG, the results of which are validated against behavioural or educational assessments.

Educational neuroscience is still an emerging area of research. It has a wide remit at present with not very well defined boundaries. A report by the Royal Society which investigated the implications of neuroscience in education (Brain Waves, 2011) states that:

> *Education is about enhancing learning, and neuroscience is about understanding the mental processes involved in learning. The common ground suggests a future in which educational practice can be transformed by science, just as medical practice was transformed by science about a century ago.* (Page v)

The aim is to understand how learning behaviour is manifested in the brain so as to improve how we practice teaching or learning or assessment activities.

Research in this area is being carried out at various institutions such as CIBSR-Stanford, the Centre for Neuroscience in Education (CNE)-Cambridge and UCL. The CNE Director, Professor Goswami, states that:

> *… the tools of cognitive neuroscience offer various possibilities to education, including the early diagnosis of special educational needs, the monitoring and comparison of the effects of different kinds of educational input on learning, and an increased understanding of individual differences in learning and the best ways to suit input to learner.* (Goswami, 2004, page 6)

The CIBSR-Stanford is also examining relationships between brain and behaviour to predict future learning difficulties in children so that early interventions could be developed.

Neuroscience has given rise to many theories about learning behaviours and classroom teaching strategies, not all of which are completely valid. Some of them may be false or incomplete or exaggerated and such misconceptions are known as 'neuromyths', a term coined by the Organisation for Economic Co-operation and Development (OECD) report on understanding the brain with

1. The National Institutes of Health (NIH) is a part of the USA Department of Health and Human Services and is the nation's medical research agency. NIH is made up of 27 Institutes and Centers, each with a specific research agenda.
2. The Big-Five model of personality assessment is a well-known model of personality (Rust and Golombok, 1999). Its five factors - Extraversion, Agreeableness, Openness to experience, Neuroticism and Conscientiousness are widely accepted to encompass most other behaviours.
3. In Figure 6, the lighter colour signifies a larger effect size, and the darker colour signifies a smaller effect.
4. The authors describe mental representation as: "the activity of neural networks of the brain which code information in the form of electrochemical activity."

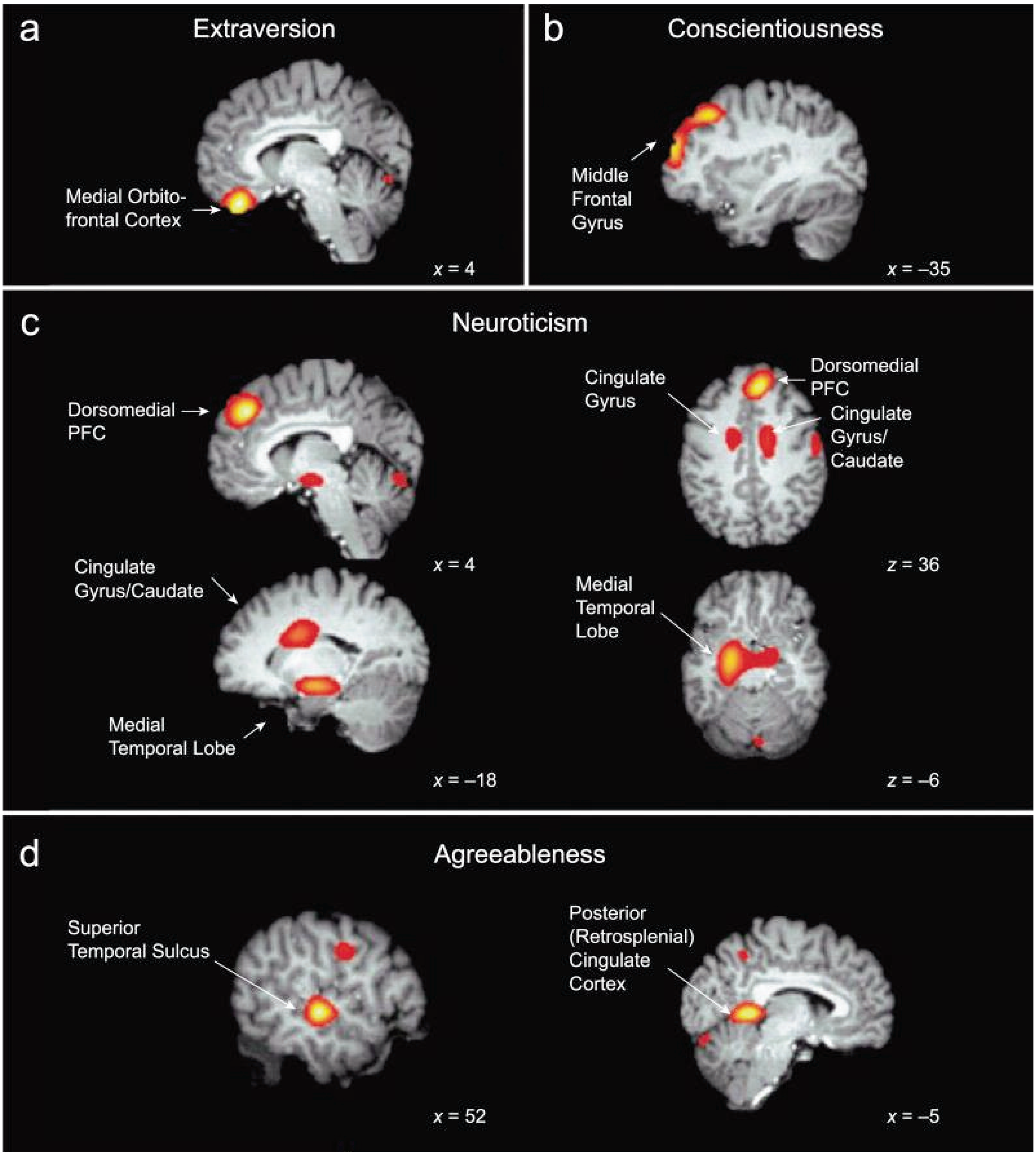


Figure 6[3]: Brain regions in which local volume was significantly associated with (a) Extraversion, (b) Conscientiousness, (c) Neuroticism, and (d) Agreeableness, as hypothesized in DeYoung *et al*. (2010). *Image courtesy of SAGE Publications.*

respect to learning (OECD, 2002). Some of the well-known myths are:

- We have either a left-brain or a right-brain thinking style
- We use only 10 per cent of our brain
- There are critical periods for learning certain tasks which cannot be learnt when that age period is over.

The myths can be difficult to dispel once they become popular. OECD (2007) points out that these misconceptions "often have their origins in some element of sound science, which makes identifying and refuting them more difficult".

A detailed review of some of the neuromyths is given in Beyerstein (1999), Geake (2008) and Goswami (2006).

An important characteristic that needs to be mentioned here is the ability of the brain to form new connections between neurons in response to new learning or environment. This characteristic is known as 'Brain Plasticity' or 'Neuroplasticity' and can occur during adult life as well. It can be defined as the ability of the brain to mould itself in response to an external sensory input or internal events which may include the effects of our own thoughts or visual imagery, hormones, genes and following brain injury (Anderson and Sala, 2012). This points to the possibility of developing new behaviours and skills later on in life and could have implications for life-long learning.

Findings from neuroscience could help in test development by giving a better insight into the underlying constructs we assess and by providing us with tools to develop better tests. We might be able to use the understanding of the brain to tap into the specific skills we intend to assess and therefore prepare tests with relatively higher validity. For instance, if brain imaging techniques can help us to establish whether a question of numerical reasoning is in fact assessing numerical reasoning, as we have defined it, and not some other skill, it can provide a significant contribution in assessing the validity of test questions. At present it appears that the most visible area of educational assessment where neuroscience is being used is the diagnosis of learning difficulties such as dyslexia and dyscalculia. The growing understanding of such learning difficulties will inevitably lead to a better understanding of skills such as language and numerical cognition which perhaps could be used in test development at some stage. An example of how neuroscience is being applied to understand dyscalculia and mathematical skills is given in the following section.

How neuroscience can help – Dyscalculia and Mathematics

This example draws heavily on Butterworth, Varma and Laurillard (2011), Butterworth and Laurillard (2010) and Szücs, Devine, Soltesz, Nobes and Gabriel (2013).

Developmental dyscalculia refers to the existence of a severe disability in learning arithmetic. It has roughly the same prevalence as dyslexia but has not received as much attention. Usually low achievement on mathematical achievement tests is used as a criterion for identifying dyscalculia. However, this approach may not necessarily identify the underlying neurological reasons and therefore may lead to insufficient remedial actions.

An understanding of how mathematical ability is represented in the brain would be very helpful for designing remedial actions for dyscalculia and for Mathematics education in general. Neuroscientists are working on understanding how mathematical skills and dyscalculia can be explained through neuro-imaging research.

One area of research (Butterworth *et al*., ibid.) suggests that dyscalculia is caused by a disorder in the way the brain represents magnitude. Neuroscientists have been able to identify areas of the brain associated with mathematical skills, such as learning new arithmetical facts (frontal lobes and the intra-parietal sulci – IPS), using previously learned facts and in retrieving facts from memory (left angular gyrus) (Ischebeck, Zamarian, Schocke and Delazer, 2009). Various experiments have shown a reduced activation in these regions for children with dyscalculia (Mussolini *et al*., 2010). The identification of the region where almost all arithmetical abilities and numerical processes are mapped (parietal lobes) can significantly help understand the basis of mathematical skills or their lack thereof. Dyscalculics show poor performance on numerosity tasks such as counting the number of dots or making number comparisons, which may suggest that dyscalculia is characterised by impairment in magnitude representation. An illustration of these two numerosity tasks is given in Figure 7.

Neuroscience experiments have shown the areas in the brain which get activated while performing such numerosity tasks. A difference in activation of these areas in normal functioning brains and those affected by dyscalculia is shown in Figure 8. The research suggests that the pedagogic interventions to help dyscalculic children should, therefore, attempt to make the individuals develop the ability to process the numerosities.

An alternate focus of research (Szücs *et al*., 2013) suggests that the magnitude representation function of the brain might not be sufficient to explain dyscalculia and that impairment of other functions such as visuo-spatial short term memory and working memory along with inhibitory functions[5] might lead to dyscalculia. They recommend that various theories, along with behavioural research, need to be tested to gain sound understanding of mathematical skills.

Teachers make an attempt to improve the performance of their students on mathematical tasks. However these students might be those having low numeracy skills and not necessarily have dyscalculia. The assessment of students is usually based on curriculum-based tests which may not necessarily differentiate between dyscalculia and general low numeracy. Neuroscience evidence, on the other hand, can provide a more targeted approach for assessment and remedial action which is largely independent of learners' social and educational circumstances (Landerl, Bevan and Butterworth, 2004; Butterworth *et al*., 2010).

The classical remedial measures require trained special education needs (SEN) teachers and considerable time, both of which are limited resources. Computer adaptive software based on neuroscience that allows learners to explore the meaning of numbers can provide an optimum solution. Examples of two such software are mentioned below.

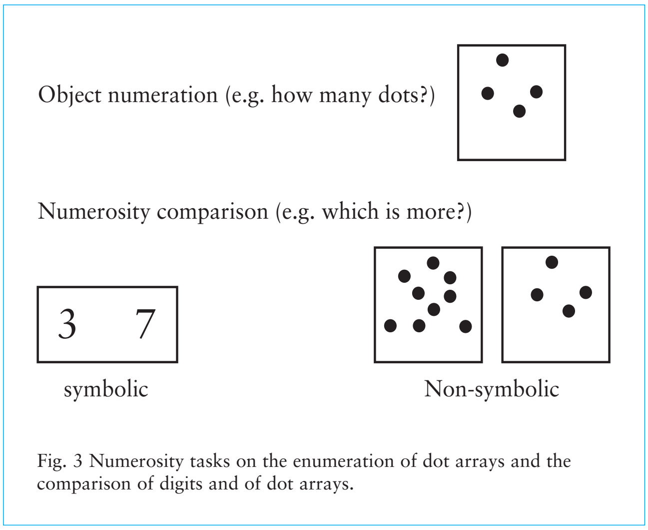


Figure 7: Example of numerosity tasks (Butterworth and Laurillard, 2010).
Image courtesy of the authors.

The *Number Race* game (Wilson, Revkin, Cohen, D., Cohen, L., and Dehaene, 2006) targets the area of the brain that supports early arithmetic to improve the precision in this skill. The learners are required to select the larger of the two arrays of dots. The software adapts to the performance of the learner, making the difference between the arrays smaller as their performance improves and provides feedback as to which is correct. Another game, *Graphogame-Maths*, targets the area of the brain known for representing and manipulating sets. Candidates are required to identify the link between the number of objects in the sets and their verbal numeric label and are given feedback about the correct answer. Some studies report some improvement in performance of children in several tasks after training which could mean an improvement in their numerical cognition (Wilson *et al*., 2006 and Räsänen, Salminen, Wilson, Aunio, and Dehaene, 2009). More such numeracy games are available from Numeracy Games (2013).

The private sector is becoming increasingly active in the area of brain-training software. According to an estimate, the market for such software

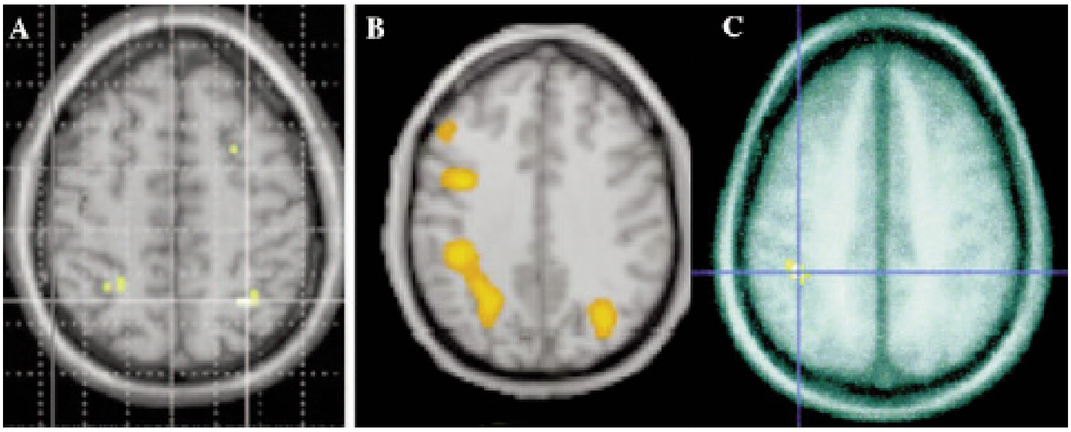


Fig. 2
Normal and atypical adult brain areas for bilateral number processing in the intraparietal sulcus.
a The *highlighted parts* show the areas that are normally activated in numerosity comparison tasks (Castelli *et al*., 2006).
b The *highlighted parts* show the networks normally activated for arithmetical calculation, which include the numerosity processing areas (Zago *et al*., 2001).
c The *highlight* indicates the part that is found to be structurally abnormal in adolescent dyscalculics (Isacs *et al*., 2001).

Figure 8: Comparison of brain activation in normal individual and those affected by dyscalculia (Butterworth and Laurillard, 2010).
Image courtesy of the authors.

in 2012 was roughly £1 billion which is expected to rise to £6 billion by 2020 (TES, 2013). However, it is important to note that such products may or may not be valid in scientific terms and might raise some ethical concerns and have unknown side effects.

Some examples were shown earlier to give an idea of how neuroscience might be used in education. Goswami (2004) notes that while some of the studies confirm what was already known from behavioural studies, new insights are also being gained such as "a way of distinguishing between different cognitive theories (e.g., whether dyslexia has a visual basis or a linguistic basis in children)." The use of neuroscience has two implications here – firstly, to complement traditional research and confirm what is already known and secondly, to give us an insight into what has hitherto been unknown. The latter holds some immensely exciting possibilities for the future.

Discussion

Knowledge of how our brains work will allow us to better understand human behaviour and cognition. Neuroscience holds the potential to enable us to provide more targeted solutions, be it in medicine or education. In recent years there has been a significant interest in applying neuroscience to various fields. After the surge of information and communication technologies, neuroscience, by allowing us to see beyond what has been observable, may well lead us to the next phase in human development history.

However, we need to be cautious of the fact that, at present, a great deal of attention being given to this field is driven by commercial reasons. The consumer marketing companies, in particular, are looking to exploit neuroscience research – not all of which may withstand scientific scrutiny. Similarly, a large number of software development companies have launched products (such as educational training and psychological tests) based on what is claimed as neuroscience evidence. It may be difficult to establish the authenticity of such products.

The techniques used for mapping brain activity (such as MRI scans) are currently very expensive. This is one of the reasons why the sample size in most neuroscience studies is small, which could affect general conclusions. The scanning machines are not very convenient for the participants.

5. Inhibitory functions refer to, for example, the ability to withhold a response or block out distracting stimuli.

They require lying down inside huge scanners or wearing caps knitted with detectors which require a long time to set up. Testing young children is even more difficult. In addition, a large number of studies are based on mapping the areas of brain which get activated when a certain task is performed. However, different areas of the brain might get activated due to different reasons (such as movement of a body part) and not necessarily the activity being monitored.

Continued research in health as well as learning disabilities will lead to applications in education as well. Some of the most important benefits of neuroscience in understanding and improving individual performance are likely to be derived from increasing understanding of how functions such as memory, attention span and reward systems work in the brain. As various teaching and learning strategies based on neuroscience start becoming popular we will have to watch out for the neuromyths which can be difficult to dislodge once they enter the popular culture. Products/methodologies based on neuroscience can be evaluated based on:

- do they provide any additional utility – over and above the current ones,
- are they scientifically reliable and valid (doing what they purport to do)

  and
- do they justify the higher costs as compared to traditional approaches?

The use of neuroscience in test development is limited by the fact that currently there isn't ample understanding of the relationship of neural signals with high level concepts such as ability and skills. The neuroscientific understanding of such concepts is still at a nascent stage. So, the development of educational tests and examination questions *purely* based on neuroscience is, perhaps, a bridge too far. However, the knowledge base in this area is expanding rapidly which could be applied in test development. For instance, the growing understanding of different strategies used by students to answer examination questions could help validate the constructs used in the tests. The most important question for a test developer is: Does an examination question measure what it is intended to measure? If neuroscience can provide an answer to this question, more accurately than what we already know, it will be a tremendous contribution to the field of assessment. As neuroscientists expand their knowledge of how different skills and behaviours are represented in the brain, we can expect neuroscience applications in educational assessment in the near future. It is also worth mentioning that test development is not a standalone activity. It is informed by various factors which are more likely to be influenced by neuroscience evidence in the immediate future than question writing itself. Such factors include teaching, curriculum, use of technology and political and social environment.

Huge investments are being made in neuroscience research in the US and Europe by governments and academic institutions. We can expect some high level research outcomes in the following years. The on-going improvement in scanning machines will also make research easier and more accessible. However, for the time being it appears that the commercial sector will continue to lead in using neuroscience.

Educational authorities and awarding organisations will need to keep themselves abreast of how developments such as neuroscience might have an impact on their operations. The prime objective of this study has therefore been to briefly encapsulate the association of neuroscience and education so as to ensure future readiness. Not having the answers now does not mean that we will not have them in the future; nor should we stop looking for them.

Acknowledgements

I would like to thank Dr Denes Szücs (Centre for Neuroscience in Education, University of Cambridge), Apoorva Bhandari (MRC Cognition and Brain Sciences Unit, Cambridge) and my former Cambridge Assessment Research Division colleague, Amy Devine (now Centre for Neuroscience in Education, University of Cambridge), for their advice.

The following individuals also provided some useful information: Prof Geraint Rees (Institute of Cognitive Neuroscience, UCL), Professor Allan L. Reiss and Shelli Kesler (Center for Interdisciplinary Brain Sciences Research, Stanford University School of Medicine) and Helen Harth (Loughborough University).

Dr Robert H. Pierzycki (National Institute for Health Research: Nottingham Hearing Biomedical Research Unit) was very helpful in explaining the working of some of the brain-scanning techniques.

All efforts were made to seek permissions for the images used in the article and I am grateful to my colleague Karen Barden for her contribution in managing the permissions and the references used. I thank the relevant sources for permitting their use and if any omissions were made, we apologise, and will be happy to make changes in the online edition of the journal.

References

ABC News (2011). *Brain Activity Measures Response to Ads, Commercials*. Available online at: http://abcnews.go.com/Technology/scientists-response-ads-measuring-brainwaves/story?id=12841570. (Accessed 10 December 2013).

Anderson, M. & Sala, S.D. (2012). Neuroscience in education: an (opinionated) introduction. In: Sala, S.D and Anderson, M. (Eds.) *Neuroscience in Education: The good, the bad, and the ugly*. Oxford University Press.

Beyerstein, B. L. (1999). Whence Cometh the Myth that We Only Use 10% of our Brains? In Sergio Della Sala. *Mind Myths: Exploring Popular Assumptions About the Mind and Brain*. Wiley. pp.3–24.

BRAIN Initiative (2013). Available online at: http://en.wikipedia.org/wiki/Brain_Research_through_Advancing_Innovative_Neurotechnologies (Accessed 10 December 2013).

Brain Waves (2011). *Brain Waves 2: Neuroscience: implications for education and lifelong learning*. The Royal Society, London. Available online at: http://royalsociety.org/policy/projects/brain-waves/education-lifelong-learning/ (Accessed 10 December 2013).

Butterworth, B. & Laurillard, D. (2010). Low numeracy and dyscalculia: identification and intervention. *ZDM Mathematics Education*, **42**, 6, 527–539. Available online at: http://www.mathematicalbrain.com/ (Accessed 17 April 2013).

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From Brain to Education. *Science*, **332**, 1049–1054. Available online at: http://www.sciencemag.org/content/332/6033/1049.short (Accessed 16 April 2013).

Center for Interdisciplinary Brain Sciences Research, Stanford School of Medicine CIBSR (2013). *About Us*. Available online at: http://cibsr.stanford.edu/about/ (Accessed 27 March 2013).

Connectome (2013). *Human Connectome Project*. Available online at: http://www.humanconnectomeproject.org/ (Accessed 25 March 2013).

Costa P.T. Jr. & McCrae R.R. (1992). *NEO PI-R professional manual*. *Psychological Assessment Resources*. Odessa, FL.

DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris X., Rajeevan, N. and Gray, J. R. (2010). Testing Predictions From Personality Neuroscience: Brain Structure and the Big Five. *Psychol Sci*. **21**, 6, 820–828. Available online at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3049165/ (Accessed 30 January 2013).

Europa (2013). *Graphene and Human Brain Project win largest research excellence award in history, as battle for sustained science funding continues*. Available online at: http://europa.eu/rapid/press-release_IP-13-54_en.htm. (Accessed 25 March 2013).

Fast Company (2011). *Neurofocus uses neuromarketing to hack your brain*. Available online at: http://www.fastcompany.com/1769238/neurofocus-uses-neuromarketing-hack-your-brain. (Accessed 30 January 2013).

Fowler, J. H. and Schreiber, D. (2008). Biology, Politics, and the Emerging Science of Human Nature. *Science* **322**, 5903, 912–914. Available online at: http://www.sciencemag.org/content/322/5903/912.full. (Accessed 16 August 2013).

Geake, J. (2008). Neuromythologies in education. *Educational Research*, **50**, 2, 123–133.

Goswami, U. (2004). Neuroscience and education. *British Journal of Educational Psychology*, **74**, 1, 1–14.

Goswami, U. (2006). Neuroscience and education: From research to practice? *Nature Neuroscience Reviews*: **7**, 5, 406–413.

Guardian (2012). *Ad men use brain scanners to probe our emotional response*. Available online at: http://www.guardian.co.uk/media/2012/jan/14/neuroscience-advertising-scanners. (Accessed 13 March 2013).

Human Brain Project (2013). *The Human Brain Project*. Available online at: http://www.humanbrainproject.eu/ (Accessed 25 March 2013).

Ischebeck, A., Zamarian, L., Schocke, M. & Delazer, M. (2009). Flexible transfer of knowledge in mental arithmetic – an fMRI study. *Neuroimage*, **44**, 3, 1103–1112.

Issenberg, S. (2012). *The Victory Lab: The Secret Science of Winning Campaigns*. Crown Publishing Group: New York.

Landerl, K., Bevan, A., Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: a study of 8–9-year-old students. *Cognition*, **93**, 2, 99–125.

Legrenzi, P., Umilta, C. & Anderson, F. (2011) *Neuromania: on the limits of brain science*. Oxford University Press: Oxford.

Magnet Lab (2013). MRI: *A Guided Tour*. Available online at: http://www.magnet.fsu.edu/education/tutorials/magnetacademy/mri/fullarticle.html. (Accessed 16 August 2013).

Mapping Techniques (2013a). *How Brain Mapping Works*. Available online at: http://science.howstuffworks.com/life/inside-the-mind/human-brain/brain-mapping1.htm. (Accessed 30 January 2013).

Mapping Techniques (2013b). Available online at: http://en.wikipedia.org/wiki/Magnetoencephalography. (Accessed 12 March 2013).

Mussolini, C., De Volder, A., Grandin, C., Schlögel, X., Nassogne, M.C., & Noël, M.P. (2010). Neural correlates of symbolic number comparison in developmental dyscalculia. *Journal of Cognitive Neuroscience*. **22**, 5, 860–74.

Numeracy Games (2013). *Developing Number Sense*. Available online at: http://www.low-numeracy.ning.com/. (Accessed 17 April 2013).

Neurocinema (2011). *Rise of Neurocinema: How Hollywood Studios harness your brainwaves to win Oscars*. Available online at: http://www.fastcompany.com/1731055/rise-neurocinema-how-hollywood-studios-harness-your-brainwaves-win-oscars (Accessed 11 December 2013).

Neurorelay (2012). *Insights from "The Buying Brain: Secrets for Selling to the Subconscious Mind" Book Review*. Available online at: http://neurorelay.com/2012/05/17/insights-from-the-buying-brain-secrets-for-selling-to-the-subconscious-mind-book-review/. (Accessed 13 March 2013).

NIH (2012). *NIH Toolbox Brochure*. Available online at: http://www.nihtoolbox.org/WhatAndWhy/Assessments/NIH%20Toolbox%20Brochure-2012.pdf. (Accessed 16 April 2013).

New York Times (2012). *Academic 'Dream Team' Helped Obama's Effort*. Available online at: http://www.nytimes.com/2012/11/13/health/dream-team-of-behavioral-scientists-advised-obama-campaign.html. (Accessed 25 March 2013).

New York Times (2013). *Obama Seeking to Boost Study of Human Brain*. Available online at: http://www.nytimes.com/2013/02/18/science/project-seeks-to-build-map-of-human-brain.html?pagewanted=all&_r=0. (Accessed 25 March 2013).

OECD (2002). *Understanding the Brain – Towards a New Learning Science.* OECD, Paris.

OECD (2007). *Understanding the Brain: the Birth of a Learning Science*. OECD, Paris.

Prism (2013). *Prism Brain Mapping: Using Neuroscience to improve performance*. Available online at: http://www.prismbrainmapping.com/ (Accessed 30 January 2013).

Räsänen, P., Salminen, J., Wilson, A.J., Aunio, P. and Dehaene, S. (2009) Computer-assisted intervention for children with low numeracy skills. *Cognitive Development*, **24**, 4, 450–472. Available online at: http://www.aboutdyscalculia.org/author.html (Accessed 9 July 2013).

Rust, J. & Golombok, S. (1999) 2nd edition *Modern Psychometrics: The Science of Psychological Assessment*, Routledge, London & New York.

Schreiber, D. (2007). Political Cognition as Social Cognition: Are we all political sophisticates? In: Neuman, W.R., Marcus, G. E., Crigler A.N., *et al*. (Eds) *The Affect Effect: Dynamics of Emotion in Political Thinking and Behavior*. Chicago, IL: University of Chicago Press, pp.48–70.

Stevenson Jr., A.E. (1952) *Speech at the University of Wisconsin*. Madison, 8 October 1952.

Szücs, D. & Goswami, U. (2007). Educational neuroscience: Defining a New Discipline for the Study of Mental Representations. *Mind, Brain and Education*, **1**(3), 114–127.

Szücs, D., Devine, A., Soltesz, F., Nobes, A., Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*. Available online at: http://www.sciencedirect.com/science/article/pii/S0010945213001688#. (Accessed 26 September 2013).

TES (2013). *Get inside their heads*. TES, 1 March 2013, pp.28–32.

Todd, R.L., Berninger, V. W., Stock, P., Altemeier, L., Trivedi, P. & Maravilla, K. R. (2011). Differences between good and poor child writers on fMRI contrasts for writing newly taught and highly practiced letter forms. *Reading and Writing*, **24**, 5, 493-516. Available online at: http://link.springer.com/article/10.1007%2Fs11145-009-9217-3 (Accessed 12 March 2013).

Westen, D. (2007). *The political brain: The role of emotion in deciding the fate of the nation*. New York. Public Affairs Books.

Wilson, A.J., Revkin, SK., Cohen, D., Cohen, L. & Dehaene, S. (2006). An open trial assessment of The Number Race, an adaptive computer game for remediation of dyscalculia. *Behavioural and Brain Functions*, **30**, 2, 20.

Book announcement: Validity in Educational and Psychological Assessment

Paul Newton Institute of Education, University of London (formerly Cambridge Assessment) and **Stuart Shaw** Cambridge International Examinations

Introduction

For almost one hundred years, divergent views on the concept of validity have proliferated. Even today, the meaning of validity is heavily contested. Despite a century of accumulated scholarship, new definitions of validity continue to be proposed, and new 'types' of validity continue to be invented (see Newton and Shaw, 2013). Yet, against the backdrop of an evolving measurement and testing landscape and the increased use of assessments across scientific, social, psychological and educational settings, validity has remained "the paramount concept in the field of testing." (Fast and Hebbler, 2004, p.i).

Validity is universally regarded as the hallmark of quality for educational and psychological measurement. But what does quality mean in this context? And to what exactly does the concept of validity actually apply? What does it mean to claim validity? And how can a claim to validity be substantiated? In a book entitled *Validity in Educational and Psychological Assessment*, due to be published in the UK by SAGE in March 2014, we explore answers to these fundamental questions.

Validity in Educational and Psychological Assessment adopts an historical perspective, providing a narrative through which to understand the evolution of validity theory from the nineteenth century to the present day. We describe the history of validity in five broad phases, mapped to the periods between:

1. the mid-1800s and 1920: gestation
2. 1921 and 1951: crystallisation
3. 1952 and 1974: fragmentation
4. 1975 and 1999: (re)unification
5. 2000 and 2012: deconstruction.

We explain how each of these phases can be characterised by different answers to the question at the heart of any validation exercise: how much and what kind of evidence and analysis is required to substantiate a claim of validity?

The book comprises six chapters. In Chapter 1 we set the scene for the historical account which follows. Chapters 2 to 5 offer readers a chronological account that delineates the phases of development of validity theory and validation practice. In Chapter 6 we propose a framework for the evaluation of testing policy, which we based on the original progressive matrix from Messick (1980).

Chapter 1: Validity and Validation

In Chapter 1 we begin by exploring a range of everyday and technical meanings of validity in order to set the scene for the historical account which follows. This is an account of validity as a technical term of educational and psychological measurement, which is important to bear in mind because the term 'validity' has very many different meanings, some of which are entirely independent of measurement. The main chapters of the book attempt to demonstrate how, even within this relatively narrow conceptualisation, its meaning is still nevertheless contested and resistant to precise definition. Yet it needs to be appreciated, from the outset, that it does mean something quite distinctive in this particular context, even if that 'something' might be difficult to articulate.

Following a discussion of the conventions used in the textbook we present an outline of the history of validity. The historical account is our attempt to describe and to explain how conceptions of validity and validation have evolved within the field of educational and psychological measurement.

Our historical account tends to focus more on concepts of validity theory than on the practice of validation. Good validation practice is the application of good validity theory. In the absence of validity theory there is nothing to guide or to defend validation practice. It is theory that constitutes the rational basis for validation practice. As we discuss each new contribution to the theory of validity, their implications in terms of a positive, operational impact upon validation practice become increasingly apparent.

Chapter 2: The Genesis of Validity (mid-1800s to 1951)

Chapter 2 covers the first two phases outlined above: a gestational period, from the mid-1800s to 1920; and a period of crystallisation, from 1921 to 1951. The chapter is heavily skewed towards the latter, as the period during which the concept of validity developed an explicit identity or, perhaps more correctly, a range of different identities.

In this chapter, we explore early conceptions of validity and validation, focusing particularly upon achievement tests, general intelligence tests, and special aptitude tests. We argue that the emergence of validity as a formal concept of educational and psychological measurement can only be understood in the context of major developments in testing for educational, clinical, occupational and experimental purposes which occurred during the second half of the nineteenth century and the early decades of the twentieth century, most notably in England, Germany, France and the USA. Upon this foundation was proposed the 'classic' definition of validity: the degree to which a test measures what it is supposed to measure.

Although there are numerous accounts of the history of validity theory and validation practice during the early years (e.g. Anastasi, 1950; Geisinger, 1992; Shepard, 1993; Kane, 2001) the impression given is often of a period almost exclusively dominated by prediction, the empirical approach to validation, and the validity coefficient. Reflecting on this period, Cronbach (1971) observed that the theory of prediction was very nearly the whole of validity theory until about 1950; a characterisation later endorsed by Brennan (2006). Kane (2001) characterised the early years as the 'criterion' phase, where the criterion was typically understood as the thing that was to be predicted.

The impression given by a number of notable chroniclers (e.g. Moss, Girard and Haniford, 2006) is that the key developments in validity theory can be traced either to successive editions of *Educational Measurement*, beginning with Lindquist (1951) or to successive editions of professional standards documents, beginning with American Psychological Association/American Educational Research Association/ National Council on Measurements Used in Education (APA, AERA, NCMUE, 1954). We argue that there is a far more interesting story to be told about the early years. We contend that many of the developments in validity theory and validation practice, from the middle of the twentieth century onwards, are simply elaborations of earlier insights. The earliest definition of validity was far more sophisticated than the idea of a validity coefficient might suggest, and the earliest approaches to validation were far more complex and involved. Education took a lead in formally defining the concept, and achievement testers, aptitude testers, intelligence testers and personality testers played their role in refining it and developing new techniques for investigating it.

The more interesting story of validity during the early years is one of sophistication and diversity; at least in terms of ideas, if not always in terms of practice. Because of its diversity, though, it is hard to characterise the period succinctly.

Chapter 3: The Fragmentation of Validity: 1952 to 1974

The diversity of ideas on validity and validation during the early years presented a challenge to test developers and publishers. Given a variety of approaches to validation to choose from, and with even the experts valuing those approaches quite differently, how were professionals in the field to decide what information on test quality they needed to make available to consumers? And, in the absence of agreement upon principles of best practice and specific guidelines about criteria for the evaluation of tests and testing practices, how were test developers and publishers to be held to account?

The first edition of what was to become known as the *Standards* (APA, AERA, NCMUE, 1954) was written to make sense of the landscape of the early years. As a consensus statement of the professions, the *Standards* included both implicit standards for thinking about validity and explicit standards for conducting and reporting validation research.

The *Standards* emphasised 'types' of validity, specialised to the contexts of test use: content validity, predictive validity, concurrent validity, and construct validity. If, for example, you needed to validate an interpretation drawn in terms of achievement, then you needed to adopt a particular approach to validation, content validation, which meant establishing a particular kind of validity, content validity. Although these were explicitly described as "Four types of validity" (APA, AERA, NCMUE, 1954, p.13) the *Standards* was a little confused over the matter, also describing them as 'aspects' of a broader conception.

Between 1954 and 1974, the *Standards* was revised twice, in order to respond to constructive criticism, to take account of progress in the science and practice of educational and psychological measurement, and to respond to societal change. Yet, mixed messages continued to be promulgated over the nature of validity. For many who were influenced by the *Standards* during this time, they came to embody and to cement a fragmented view of validity and validation, whereby different uses to which test scores were to be put implied different approaches to validation and even different kinds of validity.

Chapter 4: The (Re)Unification of Validity: 1975 to 1999

Samuel Messick's account of validity and validation became the zeitgeist of late twentieth century thinking on validity during the 1980s and 1990s. Developing ideas from Harold Gulliksen and Jane Loevinger, and with the support of allies including Robert Guion, he brought the majority of measurement professionals of his generation around to the viewpoint that all validity ought to be understood as construct validity. His thesis was that measurement ought to be understood (once more) as the foundation for all validity; and therefore that construct validation – scientific inquiry into score meaning – ought to be understood as the foundation for all validation.

Through an extended discussion of Messick's contribution to validity theory, we describe this period in terms of his *triumph* and his *tribulation*. Messick was enormously successful in promoting validity as a unitary concept, in contrast to earlier fragmented accounts. His triumph, therefore, concerned the *science* of validity: he convinced the educational and psychological measurement communities that measurement-based decision-making procedures (i.e. tests) needed to be evaluated holistically, on the basis of a scientific evaluation into score meaning. Enormously problematic, though, was his attempt to integrate values and consequences within validity theory through his famous (if not infamous) progressive matrix. Unfortunately, not only was his account confusing, it also seemed a little confused. His tribulation, it seems fair to conclude, concerned the *ethics* of validity. Messick failed to provide a convincing account of how ethical and scientific evaluation could straightforwardly be integrated.

In retrospect, it seems hard to disagree with the conclusion, drawn by Shepard (1997), that Messick's progressive matrix was a mistake. Having said that, we believe that its underlying intention was an excellent one. It was an attempt to emphasise that the following two questions were both crucial to any thorough evaluation and were inherently interrelated:

1. Is the test any good as a measure of the characteristic it purports to assess?
2. Should the test be used for its present purpose?

Messick's progressive matrix was supposed to explain the relationship between these two questions, and their relation to the concept of validity, but it was muddled. As Messick helped readers to find their way through the ambiguity of the matrix, his presentation became clearer, but also narrower, as scientific questions of test score meaning began to gain prominence while ethical questions of test score use were nudged into the wings.

Unfortunately, Messick's tribulation led to one of the most notorious debates of all time concerning the scope of validity theory. The field is now genuinely split as to whether, and if so how, evidence from consequences ought to be considered part of validity theory - an issue we tackle in Chapter 5.

Chapter 5: The Deconstruction of Validity (2000 to 2012)

During the 1990s, work on validity and validation was heavily influenced by Messick. The fifth edition of the *Standards* (American Educational Research Association/ American Psychological Association/National Council on Measurement in Education, 1999) was essentially a consensus interpretation of his position, that is, a unified conception of validity. The *Standards* reflected the prevailing view of the time - a construct-centred approach to validity. Yet, with the turn of the millennium, cracks began to emerge. On one hand, it was unclear how to translate construct validity theory into validation practice. On the other hand, it was unclear whether construct validity was actually the best way to unify validity theory. It seemed that an element of deconstruction might be in order, reflecting the desire to simplify validation practice as well as the desire to simplify validity theory.

In terms of validation practice, this period was characterised by growing consensus over the value of a new methodology for guiding, and simplifying, validation practice. Argumentation, it now seemed, held the key. Michael Kane had developed a methodology to support validation practice, grounded in argumentation (e.g. Kane, 1992). This provided a framework, or scaffold, for constructing and defending validity claims. Thus, while Messick defined the claim to validity in terms of an overall evaluative judgement, Kane explained exactly how that claim to validity could be constructed and defended. The argument-based approach took a long time to take root, though, and only began to have a significant impact well into the new millennium. In fact, even having begun to take root, it still proved surprisingly challenging to implement. Goldstein and Behuniak (2011) noted that very few examples are available to the research community of validity arguments for large-scale educational assessments.

In terms of validity theory, this period was characterised by growing controversy, embodied in two major debates. The first concerned the nature and significance of construct validity: a debate over the relatively narrow, scientific issue of score meaning. A critical question was whether construct validity ought to be considered the foundation of all validity, as Messick had argued. Related questions concerned whether all validation needed to be understood in terms of constructs; whether the nomological networks of Cronbach and Meehl (1955) were useful or even relevant to validation; whether validity was a concept more like truth or more like justified belief; whether validity ought to be theorised in terms of measurement; and whether the concept of validity could be applied in the absence of standardised procedures.

The second concerned the scope of validity: a debate over whether the concept ought to be expanded beyond the relatively narrow, scientific issue of score meaning, to embrace broader ethical issues concerning the consequences of testing. Various 'camps' developed: from liberals, who extended the use of 'validity' to embrace social considerations of test score use; to conservatives, who restricted the use of 'validity' to technical considerations of test score meaning.

Chapter 6: 21st Century Evaluation

The concept of validity has assumed a pivotal role across decades of debate on the characteristics of quality in educational and psychological measurement. Despite this, it has proved extremely resistant to definition. In Chapter 6, we respond to the concerns of the more conservatively minded, who object that the concept of validity is becoming so large as to present an obstacle to validation practice. We do so by proposing a new framework for the evaluation of testing policy. In fact, we see this as a revision of the original progressive matrix from Messick (1980), which we have redesigned to dispel some of the confusion engendered by its original presentation. After first defending the new framework we then provide a more detailed analysis of technical and social evaluation, before considering evaluation within each of the cells respectively.

Validity in Educational and Psychological Assessment will be available from March 2014. The authors believe that this book will be of interest to anyone with a professional or academic interest in evaluating the quality of educational or psychological assessments, measurements and diagnoses.

ISBNs: Paperback: 9781446253236 Hardback: 9781446253229

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.

American Psychological Association, American Educational Research Association, and National Council on Measurements Used in Education. (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin*, **51**, 2, Supplement.

Anastasi, A. (1950). The concept of validity in the interpretation of test scores. *Educational & Psychological Measurement*, **10**, 1, 67–78.

Brennan, R. L. (Ed.). (2006). *Educational Measurement* (4th ed.). Westport, CT: Praeger Publishers.

Cronbach, L.J. (1971). Test validation. In R.L. Thorndike (Ed.). *Educational Measurement* (2nd Edition) (pp.443–507). American Council on Education. Washington: DC.

Cronbach, L.J. & Meehl, P.E (1955) Construct validity in psychological tests. *Psychological Bulletin*, **52**, 4, 281–302.

Fast, E.F. & Hebbler, S. with ASR-CAS Joint Study Group on Validity in Accountability Systems. (2004). *A Framework for Examining Validity in State Accountability Systems*. Washington, DC: Council of Chief State School Officers.

Geisinger, K.F. (1992). The metamorphosis of test validation. *Educational Psychologist*, **27**, 2, 197–222.

Goldstein, J. & Behuniak, P. (2011). Assumptions in alternate assessment: An argument-based approach to validation. *Assessment for Effective Intervention*, **36**, 3, 179–191.

Kane, M.T. (1992). An argument-based approach to validity. *Psychological Bulletin*, **112**, 3, 527–535.

Kane, M.T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, **38**, 4, 319–342.

Lindquist, E.F. (Ed.) (1951). *Educational Measurement*. Washington, DC: American Council on Education.

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, **35**, 11, 1012–1027.

Moss, P. A., Girard, B. J & Haniford, L. C. (2006). Validity in Educational Assessment. *Review of Research in Education*, **30**, 1, 109–162.

Newton, P.E. & Shaw, S.D. (2013). Standards for talking and thinking about validity. *Psychological Methods*, **18**, 3, 301–319.

Shepard, L.A. (1993). Evaluating test validity. *Review of Research in Education*, **19**, 405–450.

Shepard, L.A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, **16**, 2, 5–8.

Research News

Karen Barden Research Division

Conferences and seminars

The Future of Education International Conference

In June, Sanjana Mehta attended The Future of Education Conference in Florence, Italy. The conference aims to promote transnational cooperation and share good practice in the field of innovation for education. Sanjana presented a paper on *Thrown in at the deep end? Exploring students', lecturers' and teachers' views on additional support lessons at university*.

The Assessment in Higher Education Conference

Held in Birmingham in June, this fourth biennial conference provided an opportunity to debate key issues and developments in current assessment, policy and practice. Simon Child presented a paper entitled *"I've never done one of these before". A comparison of the assessment 'diet' at A level and the first year of university*.

British Education Studies Association (BESA)

The ninth BESA Annual Conference took place at Swansea Metropolitan University in June. The key theme of the conference was Education: Past, Present and Future. Jackie Greatorex presented on *Using scales of cognitive demand in a validation study of Cambridge International A and AS level Economics*.

Journal of Vocational Education and Training (JVET)

The JVET tenth international conference was held in July at Worcester College, Oxford. Colleagues from the Research Division presented the following papers:

Jackie Greatorex: *How can major research findings about returns to qualifications illuminate the comparability of qualifications?*

Martin Johnson: *Insights into contextualised learning: how does feedback on performance contribute to professional examiners' shared understanding?*

Carmen Vidal Rodeiro: *Do Cambridge Nationals support progression to further studies at school or college, to higher education courses and to work-based learning?*

International Meeting of the Psychometric Society (IMPS)

Tom Benton presented a paper at the 78th Annual Meeting of the Psychometric Society in Arnhem, The Netherlands, in July on *An empirical assessment of the Guttman's Lambda 4 reliability coefficient*.

European Association for Research on Learning and Instruction (EARLI)

Tom Bramley presented a paper at the 15th Biennial EARLI Conference held in Munich, Germany on *Maintaining standards in public examinations: why it is impossible to please everyone*.

British Educational Research Association (BERA)

The BERA Annual Conference was held from 3–5 September at the University of Sussex. Colleagues from the Research Division and OCR presented the following papers:

Tom Benton: *Examining the impact of tiered examinations on the aspirations of young people*.

Jackie Greatorex: *How can major research findings about returns to qualifications illuminate the comparability of qualifications?*

Simon Child, Sanjana Mehta, Frances Wilson, Irenka Suto and Sally Brown: *Beyond freshers' week: An exploration of universities' assessment practices and additional support lessons*.

Sanjana Mehta, Martin Johnson, Nicky Ruston and Simon Child: *Controlled Assessment and Modern Foreign Language (MFL) speaking and listening: using a mixed methods approach to evaluate the effects of assessment arrangements*.

Magda Werno: *Addressing the needs of non-native speakers entering English education at the secondary stage*.

Tom Benton of the Research Division won the BERA poster competition with his entry entitled *Calculating the reliability of complex qualifications*. Tom was helped with the poster design by Research Division colleague Jo Ireland.

The following posters were also presented:

Nicky Rushton: *Changing times, Changing qualifications*.

Jackie Greatorex, Stuart Shaw, Phineas Hodson and Jo Ireland: *Do the A and AS level Economics examination papers elicit responses that reflect the intended construct?*

European Conference on Educational Research (ECER)

In September, Frances Wilson and Nicky Rushton attended the ECER conference in Istanbul, Turkey. The main theme of the conference was *Creativity and Innovation in Educational Research*. Frances presented a paper entitled *A comparison of assessment at school and university: More than just increasing demand*. Nicky gave a poster presentation on *The Register of Change*.

European Conference on Computer Supported Cooperative Work (ECSCW)

Martin Johnson presented a paper at the 13th European Conference on Computer Supported Cooperative Work in Paphos, Cyprus in September on *'Seeing what they say': mapping the characteristics of effective remote feedback*.

European Association of Test Publishers (E-ATP)

The E-ATP Annual Conference took place in September in St Julian's, Malta. One of the key themes was *Innovations in Assessment* and Irenka Suto presented a paper entitled *The Cambridge Approach to 21st Century skills: advances in teaching, assessment and support for learners*.

Association for Educational Assessment – Europe (AEA-Europe)

The AEA-Europe Annual Conference took place in Paris in November with the theme of *International surveys, policy borrowing and national assessment*. Several colleagues from Cambridge Assessment attended the conference and the following papers were presented:

Tom Benton and Tim Gill: *Investigating the relationship between aspects of countries' assessment systems and achievement in PISA*.

Victoria Crisp: *The judgement processes involved in the moderation of teacher-assessed projects in a national assessment*.

Gill Elliott: *Method in our madness? The advantages and limitations of mapping other jurisdictions' educational policy and practice*.

Tim Oates: *Why Grand Theory and detailed narrative are equally essential in drawing from transnational comparisons*.

Stuart Shaw, Martin Johnson and Paul Warwick: *Assessment for learning in international contexts: approaches and challenges in researching teacher values and practices (Phase 2)*.

Further information on all conference papers can be found on the Cambridge Assessment website: http://www.cambridgeassessment.org.uk/our-research/all-published-resources/conference-papers/

Publications

The following articles have been published since Issue 16 of *Research Matters:*

Crisp, V. & Grayson, R. (2011) Modelling question difficulty in an A level Physics examination. *Research Papers in Education: Policy and Practice*, **28**, 3, 346–372.

Greatorex, J. and Shaw, S. D. (2013). The validity of tutor assessed Independent Research Reports contributing to a pre-university qualification *Educationalfutures*, BESA, **6**, 1, 58–78. Available online at: http://educationstudies.org.uk/?post_type=journal&p=1113

Johnson, M (2013). 'Seeing what they say': mapping the characteristics of effective remote feedback, In: Korn, M., Colombino, T., and Lewkowicz, M. (Eds) *The 13th European Conference on Computer Supported Cooperative Work Adjunct Proceedings*. Aarhus: Aarhus University, pp.141–147. Available online at: http://ojs.statsbiblioteket.dk/index.php/daimipb/article/view/13587/11586

Keating, A. & Benton, T. (2013). Creating cohesive citizens in England? Exploring the role of diversity, deprivation and democratic climate at school. Education, *Citizenship and Social Justice*, **8**, 2, 165–184. Available online at: http://esj.sagepub.com/content/8/2/165.abstract

Mehta, S., Child, S., Brown, S. & Suto, I. (2013). Thrown in at the deep end? Exploring students', lecturers' and teachers' views on additional support lessons at university. *Future of Education Conference Proceedings 2013*. Available online at: http://conference.pixel-online.net/foe2013/common/download/Paper_pdf/132-SOE38-FP-Mehta-FOE2013.pdf

Shaw, S. D. and Imam, H. (2013) Assessment of International Students Through the Medium of English: Ensuring Validity and Fairness in Content-Based Examinations. *Language Assessment Quarterly*, **10**, 4, 452–475.

Statistical Reports

The Research Division

The on-going 'Statistics Reports Series' provides statistical summaries of various aspects of the English examination system such as trends in pupil uptake and attainment, qualifications choice, subject combinations and subject provision at school. These reports, mainly produced using national-level examination data, are available on the Cambridge Assessment website: http://www.cambridgeassessment.org.uk/ca/Our_Services/Research/Statistical_Reports.

The most recent additions to this series are:

- Statistics Report Series No.52: Popularity of A level subjects among UK university students.
- Statistics Report Series No.53: Candidates awarded the A* grade at A level in 2012.
- Statistics Report Series No.54: Provision of GCE A level subjects 2012.
- Statistics Report Series No.55: Uptake of GCE A level subjects 2012.
- Statistics Report Series No.56: Provision of GCSE subjects 2012.
- Statistics Report Series No.57: Uptake of GCSE subjects 2012.
- Statistics Report Series No.58: Provision of level 3 qualifications in English schools 2008-2012.
- Statistics Report Series No.59: Uptake of level 3 qualifications in English schools 2008-2012.
- Statistics Report Series No.60: Provision of level 2 qualifications in English schools 2008-2012.
- Statistics Report Series No.61: Uptake of level 2 qualifications in English schools 2008-2012.
- Statistics Report Series No.62: Provision of IGCSE subjects 2012.
- Statistics Report Series No.63: Uptake of IGCSE subjects 2012.
- Statistics Report Series No.64: The accuracy of forecast grades for OCR A levels in June 2012.
- Statistics Report Series No.65: A level Uptake and Results, by Gender 2003-2012.
- Statistics Report Series No.66: GCSE Uptake and Results, by Gender 2003-2012.
- Statistics Report Series No.67: A level Uptake and Results, by School Type 2003-2012.
- Statistics Report Series No.68: GCSE Uptake and Results, by School Type 2003-2012.

CONTENTS : Issue 17 January 2014

Cambridge Assessment
1 Hills Road
Cambridge
CB1 2EU

Tel: 01223 552666
Fax: 01223 552700
Email: ResearchProgrammes@cambridgeassessment.org.uk
http://www.cambridgeassessment.org.uk

ISSN: 1755–6031