

QUANTITATIVE LITERACY OF SCHOOL LEAVERS ASPIRING TO HIGHER EDUCATION IN SOUTH AFRICA

V. Frith*

Academic Development Programme

e-mail: Vera.frith@uct.ac.za

R. Prince*

Centre for Educational Testing and Placement

e-mail: Robert.prince@uct.ac.za

*University of Cape Town

Rondebosch, Cape Town

ABSTRACT

It is important for higher education educators to understand the quantitative literacy (QL) competencies of incoming students, in order to make appropriate assumptions about prior knowledge and to design suitable curricula. In this article we analyse the results of a National Benchmark Tests project's (NBTP) QL test written by a large cohort of prospective applicants to higher education, in order to contribute to this understanding. A large proportion of these candidates were unable to cope with quantitative literacy demands of the kinds commonly encountered in higher education. More than half of candidates need some kind of supplementary QL support and at least 30 per cent require extensive support. Candidates' performance on subgroups of the QL test items and on individual items provides further insights into particular strengths and weaknesses. The results highlight that opportunities for the development of QL competencies should be found and these should be integrated into disciplinary teaching.

Keywords: quantitative literacy, National Benchmarks Tests, university applicants, diagnostic assessment, curriculum, quantitative literacy proficiency

INTRODUCTION

South Africa needs to produce increasing numbers of good quality graduates, but higher education's 'graduate output has been found to have major shortcomings in terms of overall numbers, equity and the proportion of the student body that succeeds' (Council on Higher Education 2013, 15). It is common knowledge that the South African school system does not adequately prepare many students for higher education and that there is great inequality in the conditions and experiences of those who do enter the South African higher education (HE) sector. It has been pointed out by Scott, Yeld and Hendry (2007, 42) that '... the educational

factor to which poor performance is perhaps most commonly ascribed across the higher education sector is student underpreparedness for standard undergraduate programmes'. This 'underpreparedness' arises not only from content knowledge in the various school subjects that may not have been adequately taught and learned, even more crucially it arises from deficiencies in the area of academic literacies. 'The real key to whether a student will pass or fail relates to the literacy practices she brings with her to the University from her school and home environments, and the extent to which these have commonalities with the literacy practices of her chosen discipline' (McKenna 2009, 8). One of the academic literacies in which students experience particular difficulties is quantitative literacy (QL), also known as numeracy. This can be seen in results from the NBTP. For example, in the pilot tests in 2009 only one quarter of university students tested was classified as 'Proficient' in quantitative literacy (Yeld 2009, 79). In 2014, 76 693 candidates wrote the NBTP QL test as prospective applicants to higher education in 2015. In this case only 11 per cent of the candidates were classified as 'Proficient' and the remaining 89 per cent of candidates were expected to experience academic challenges due to their low levels of QL proficiency. Nearly 40 per cent were classified as 'Basic' meaning that they would experience severe academic challenges in higher education (Centre for Educational Testing for Access and Placement 2015, 26).

Critical theorists such as Boughey (2009, 4) stress that the notion of underpreparedness implies deficiency in the students only and does not recognise that higher education institutions themselves are underprepared to meet the needs of the students that they admit. University teaching needs to take into account the capabilities of the students that they accept and make changes to the curriculum to address the 'articulation gap' (Scott, Yeld and Hendry 2007, 42) between the demands of curricula and the level of many students' quantitative (and other) literacies. In order to design a more responsive curriculum, lecturers and curriculum developers in higher education need information about the capabilities of students. Providing this kind of diagnostic information is one of the goals of the NBTP: 'To inform the nature of foundation courses and curriculum responsiveness' (Griesel 2006, 4).

In this article we report on diagnostic information from the NBTP quantitative literacy test in order to contribute to this process of designing responsive curricula. While the importance of quantitative literacy for higher education is widely recognised (see, for example, Steen 2004) there is also an increasing awareness that many academic disciplines make complex quantitative demands that are often very different from those that are the focus of traditional mathematics courses. We will elaborate further on the concept of QL when we discuss the theory underlying the NBTP quantitative literacy test construct below.

We provide an overview of the results of the NBTP QL test written by a large sample of successful school leavers across South Africa who were intending to apply to enter higher education institutions in 2012. It is not our intention to provide an in-depth analysis of the performance of different categories of candidates or to analyse the factors that contribute to the performance reported. We offer this information to alert higher education lecturers to the fact that they cannot assume a high level of quantitative literacy in their first year students and to point to areas where quantitative literacy development will have to be addressed in HE curricula. It also serves to inform school teachers of the requirements of higher education in the area of quantitative literacy.

Candidates writing the National Benchmark tests (NBTs) sign a declaration allowing the use of their scores for research purposes and they are simultaneously assured of anonymity in such use. The research reported in this article has been subjected to ethical review and approved by the research ethics committee of the Faculty of Higher Education Development at UCT.

THE NATIONAL BENCHMARK TESTS PROJECT IN QUANTITATIVE LITERACY

The aims of the National Benchmark Tests Project

The NBTP was commissioned in 2005 by Higher Education South Africa (HESA). The main objective of the project is to assess the entry-level academic proficiency of all prospective students in Academic Literacy (AL) and QL, and, for candidates planning to enter courses or programmes which have a significant mathematical component, proficiency in Mathematics (MAT). In addition, the project was also conceptualised to provide a service to higher education institutions requiring information additional to that provided by the national school-leaving examination to assist in the selection and placement of students in appropriate curricular routes. The project also assists with curriculum development through first-year teaching forums with an emphasis on foundation programmes and other support courses.

The NBTs are designed specifically to complement the South African school-leaving qualification, the National Senior Certificate (NSC). They aim to provide independent and objective information against which the performance of students on the NSC can be compared and calibrated. They assess students' levels of academic readiness prior to possible entry to higher education, with the aim of providing information to facilitate appropriate placement of students in regular, augmented or foundation programmes. The tests assess competence in the key areas of academic literacy, quantitative literacy and mathematics, which represent core areas of competency in which students entering any form of higher education would be expected

to display minimum levels of proficiency. These tests are criterion-referenced, which means that they are ‘constructed to provide information about the level of a test-taker’s performance in relation to clearly defined domains of content and/or behaviours (e.g. reading, writing, mathematics) that require mastery’ (Foxcroft 2006, 9). A standard-setting process defines minimum (benchmark) scores that represent the levels at which a student would be expected to perform in order to be deemed ‘recommendable’ for different forms of educational provision.

The NBTP QL test aims to address a single, overarching question, namely ‘What are the Quantitative Literacy levels of proficiency of the school-leaving population who wish to continue with higher education?’ This is assessed at a point prior to their entry into higher education at which they could realistically be expected to cope with the QL demands of higher education study. The construct of the QL test outlined in Frith and Prince (2006, 28) is based on testing this question. Benchmark levels, based on candidates’ test scores, have been set in order to define levels of proficiency which provide guidance to the most appropriate forms of QL provision for those candidates.

Theoretical framework for the NBTP Quantitative Literacy test

The construct of the test is based on the following theoretical considerations about the nature of quantitative literacy. The nature and definition of Quantitative Literacy are actively debated, particularly in Australia and England (where it is usually called ‘numeracy’) and in the United States (where it is most often called ‘quantitative literacy’). This debate concerns itself not only with the definition of the concept, but also with its relationship to mathematics itself. Hughes-Hallet (2001, 94) expresses the distinction between quantitative literacy and mathematics as follows: ‘Mathematics focuses on climbing the ladder of abstraction while quantitative literacy clings to context. ... Mathematics is about general principles that can be applied in a range of contexts; quantitative literacy is about seeing every context through a quantitative lens.’

This idea that quantitative literacy is mainly concerned with mathematics and statistics used in context is fundamental to all its definitions, whether it is seen as social practice (Street 2005; Street and Baker 2006), a critical approach (Johnston 2007) or a component of a more general idea of literacy (Chapman and Lee 1990). ‘At the very least then, the definitions garnered from this debate would agree that numeracy is to do with “using maths in context” and that to be numerate is to have the “capacity to use maths effectively in context”.’ (Johnston 2002, 4).

In South Africa, for the Further Education and Training curriculum, the definition of Mathematical Literacy is ‘a subject driven by life-related applications of mathematics’

(Department of Education 2003, 9). If we think about quantitative literacy in higher education from this perspective, we focus on those ‘life-related’ applications that occur within the ‘life’ of academic disciplines. It is clear, regardless of one’s preferred definition, that quantitative literacy must not be seen merely as a set of generic mathematical skills and techniques. For this reason almost all items in the NBTP QL test require the practice of quantitative thinking within a given realistic context. The definition of quantitative literacy that underpins the NBTP QL test is as follows:

Quantitative literacy is the ability to manage situations or solve problems in practice, and involves responding to quantitative (mathematical and statistical) information that may be presented verbally, graphically, in tabular or symbolic form; it requires the activation of a range of enabling knowledge, behaviours and processes and it can be observed when it is expressed in the form of a communication, in written, oral or visual mode. (Frith and Prince 2006, 30).

The development of this definition was most strongly influenced by the definition of numerate behaviour underlying the assessment of numeracy in the Adult Literacy and Lifeskills (ALL) Survey (Gal, Van Groenestijn et al. 2005, 152) and the New Literacies Studies’ view of literacy as social practice (Street 2005; Street and Baker 2006; Kelly, Johnston and Baynham 2007). In higher education there are different quantitative literacy practices associated with different academic disciplines and these practices are often implicit. The curriculum exposes students to these practices and expects them to become competent practitioners. Since the NBTP QL test is designed for all higher education applicants, and must have face validity for all disciplines, it is necessary to include items that assess a range of different kinds of quantitative literacy competencies in a wide range of contexts.

The test construct is based on the idea that QL can be described in terms of 1) the contexts that require the activation of quantitative literacy practice; 2) the mathematical and statistical content that is required when quantitative literacy is practised; and 3) the underlying reasoning and behaviours (competencies) that are called upon to respond to a situation requiring the activation of quantitative literacy practice. In the test construct each item is characterised according to three dimensions: the competencies it calls for, the main mathematical and statistical ideas it employs and the level of cognitive complexity it demands. Items are not classified according to the context they are situated in, but tests include a range of different kinds of contexts. This is important because familiarity or unfamiliarity with a particular context for a particular candidate will most likely affect its level of difficulty for that candidate.

The NBTP QL test in practical terms assesses a candidate’s ability to:

- Understand and use a range of quantitative terms and phrases.

- Competently interpret quantitative information.
- Read and interpret tables, graphs, charts, diagrams and texts and integrate information from different sources.
- Apply quantitative procedures in various situations and do simple calculations and estimations which may involve multiple steps.
- Formulate and apply simple formulae.
- Identify trends and patterns in various situations.
- Interpret two-dimensional representations of three-dimensional structures.
- Reason logically.

The structure, administration and scoring of the NBTP QL test

The NBTP QL test comprises fifty multiple choice items, selected according to the QL domain-specific specification table (Frith and Prince 2006, 32), which classifies each item in terms of the competencies it addresses and the mathematical and statistical ideas it draws on. A summary of the QL domain is provided above. The items in the test do not assume knowledge of the content of any specific school leaving subject but rather assume the generic competencies acquired through schooling.

The NBTs are administered at test centres under standardised conditions. Invigilation training for the tests, including the QL test, is accomplished in a one-day session in which invigilators are familiarized with the test instructions and the general procedures to be adhered to during the test administration. During training, invigilators are provided with the 'National Benchmark Tests Project: Operational Manual' which includes the standardized instructions to be used in the actual testing session.

The QL test is administered as two out of seven sections of a larger Academic and Quantitative Literacy (AQL) test in paper and pencil format. Candidates are provided with a test book, a pencil and an answer sheet. Candidates are given 30 minutes to complete each section of the QL test.

Candidates' responses are recorded on mark-reading (bubble answer) sheets that are scanned using Optical Scanner technology. Responses are scored using the unidimensional three-parameter (a, b, c) Item Response Theory (IRT) model, where a = discrimination, b = difficulty, and c = guessing/pseudo-chance (Yen and Fitzpatrick 2006, 114).

QL items are scored dichotomously, i.e. either as right or wrong. Since all tests are power tests, the first missing response in a section is scored as wrong and later ones in a section as 'not reached', which means that these items do not contribute to a candidate's overall score.

This is valid, given that piloting and the experience of several years shows that sufficient time has been allocated to each of the sections in the QL test.

Scores on different versions of the QL test are also linked and equated (Holland and Dorans 2006) to ensure that performance on different versions of the test is comparable and is not a function of the version of the test that the candidate has written.

The benchmarks for the NBTP QL test

The NBTs are criterion-referenced tests and the NBTP aims to deliver information against benchmarked categories of performance for formal study at institutions of higher learning. Table 1 provides a description of the QL benchmark levels defining proficiency levels for degree study, the score ranges and what institutional responses to candidates performing at these levels should be.

The score range endpoints at which the benchmarks are defined were first set in May 2009 (after a pilot exercise in 2008) by panels drawn from across the country, comprising academics who were at that stage engaged in mainstream teaching relevant to the domain and who had not previously been involved in any NBTP test development processes. Benchmarks were revised in 2012 as part of good testing practice. The standards-setting workshops were led by a senior psychometrician from the Educational Testing Service (ETS), Princeton, New Jersey. These benchmarks were set using the ‘modified Angoff’ method (Hambleton and Pitoniak 2006).

Table 1: Current NBTP QL test benchmarks for degree study set in 2012

Performance band	Score Range	Description
Proficient	70%-100%	Test performance suggests that future academic performance will not be adversely affected (students may pass or fail at university, but this is highly unlikely to be attributable to strengths or weaknesses in the domains tested). If admitted, students may be placed into regular programmes of study.
Intermediate	38%-69%	The challenges identified are such that it is predicted that academic progress will be adversely affected. If admitted, students' educational needs should be met as deemed appropriate by the institution (e.g. extended or augmented programmes, special skills provision).
Basic	0%-37%	Test performance reveals serious learning challenges: it is predicted that students will not cope with degree-level study without extensive and long-term support, perhaps best provided through bridging programmes (i.e. non-credit preparatory courses, special skills provision) or FET provision. Institutions admitting students performing at this level would need to provide such support themselves.

Given that the Intermediate band comprises the majority of the applicant pool, it is students in this band for whom educational institutions need guidance on the most appropriate forms of support that should be provided to address their educational needs. Thus the Intermediate performance band has been divided into Upper and Lower Intermediate as shown in Table 2.

These bands were not set through the standards-setting workshops but the division of the Intermediate band has been found to be very effective in differentiating the levels of support that would be most appropriate for students with scores in this band.

Table 2: NBTP QL test degree Intermediate benchmarks and how they should be interpreted

Intermediate performance band	Score Range	Description
Intermediate Upper	54%-69%	Students are likely to need complementary support (additional tutorials, workshops, augmented courses, language intensive work)
Intermediate Lower	38%-53%	Students need to be placed in an extended programme

NBTP QL test item subgroup analysis

For the purposes of analysis and reporting, items in the test were categorised according to two different dimensions, based on two different dimensions of the test construct: firstly, the main mathematical and statistical ideas that the item addresses and secondly, the main competence that a candidate must exercise in order to answer the item correctly. This was done so that performance on each of the categories could be reported separately in order to provide more detailed diagnostic information about candidates' competencies.

Within the 'main mathematical and statistical ideas' dimension each item was assigned to one of the following six categories:

- Quantity, number and operations. These items focus on understanding the decimal number system, orders of magnitude, absolute and relative quantities, ratios, operations, simple calculation and estimation.
- Data representation and analysis. Examples are items that require candidates to read off and interpret information from charts, graphs and tables as well as items that require a basic understanding of data distributions and descriptive statistics.
- Shape, dimension and space. Items in this subgroup focus mainly on basic geometrical knowledge and familiarity with units of measurement, interpretation of scale diagrams and the ability to visualise simple three-dimensional structures.
- Relationships, pattern, permutation. These items focus on translating between verbal, graphical and tabular representations of relationships and on recognising number patterns.
- Change and rates. These items deal with comparing numbers, describing change in absolute and relative terms and rates of change represented graphically.
- Chance and uncertainty. Items in this subgroup test the understanding of the concept of chance and the ways that it is represented numerically. There are very few items in this category and for this reason scores on this subgroup are not reported in this article.

Items were also each assigned to one of six categories on the basis of the main competency which the items were designed to assess. These were as follows:

- Reasoning and synthesising. An example is an item that requires the candidate to identify and read off appropriate values from two different charts in order to perform a calculation. Another example is the one described under the heading ‘Proportional reasoning’ below.
- Translating between different representations. An example is an item that gives a verbal description of a relationship and requires a candidate to identify the graph representing that relationship.
- Interpreting problem statements and computing. These are items where a calculation is required, for example, calculating the area of the walls of a room with given dimensions or using a given formula to calculate an area. (Note that calculators are not used by candidates, but all calculations are simple – for example, fractions can readily be simplified by cancellation, numbers used have easily identifiable factors, etc.)
- Deriving information from data representations. These items require one or more values to be read off charts or tables, for example, to identify where two different quantities are in a given ratio or to say how long it would take for a quantity to be increased by a given factor.
- Knowing simple facts and applying them. An example is an item for which the correct answer depends on knowing the fact that the mean and the median of a symmetrical distribution are the same.
- Extrapolating and visualising. These items are ones where the candidate has to recognise patterns or predict terms in a sequence. Other items in this category require visualisation in three dimensions.

RESULTS

The results are presented for a large sample ($n = 6\,326$) of school-leavers from across South Africa who in 2011 wrote one version of the NBTP QL test. After showing some details about the sample we present the overall distribution of scores for the whole sample on the entire QL test as well as the distributions for males and females separately. In addition, we report on the distributions of the performances of the whole sample on subgroups of items in the test, that is, collections of items defined in terms of the main mathematical and statistical ideas and in terms of the competencies they assess. Finally, we discuss the performance of the candidates in the different proficiency bands on some examples of individual items, not only in terms of the

proportions who answered correctly, but also the proportions who selected certain alternative answers.

The characteristics of the sample of test candidates

Some self-reported demographic characteristics of this sample are shown in Table 3. The majority of candidates were African and did not speak English as their home language. English is however the most frequently reported home language, spoken by about one third of candidates. There were more female candidates than males in this sample, which is consistent with national cohorts of NBTP test candidates.

Table 3: Demographic characteristics of the 6 326 NBTP QL test candidates

		Number	Percentage (%)
Gender	Male	2 530	40.0
	Female	3 760	59.4
	Not specified	36	0.06
Population group	African	3 995	63.2
	Coloured	433	6.8
	Indian	527	8.3
	White	1 312	20.7
	Not specified	59	1.0
Home language	Afrikaans	264	4.2
	English	2 209	34.9
	isiNdebele	50	0.8
	isiXhosa	981	15.5
	isiZulu	732	11.6
	Sesotho	430	6.8
	Sesotho sa Leboa	464	7.3
	Setswana	323	5.1
	siSwati	245	3.9
	Tshivenda	206	3.3
	Xitsonga	205	3.2
	Other	217	3.4

It must be stressed that the candidates in this sample were school leavers intending to apply to higher education. It is not possible to say what proportion of these candidates (and which of them) actually became higher education students and so it is not possible to say to what extent this sample is representative of actual first year students. However, it is representative of the pool of school leavers from which higher education draws its students.

Distributions of scores for entire test

Figure 1 shows how the scores for the entire sample of candidates were distributed among the proficiency bands. It is clear that most were in the lower bands and fewer than 20 per cent were in the Proficient band. This suggests that the majority of candidates aiming to enter higher education are in need of some kind of supplementary support while nearly one third of the candidates would require extensive support.

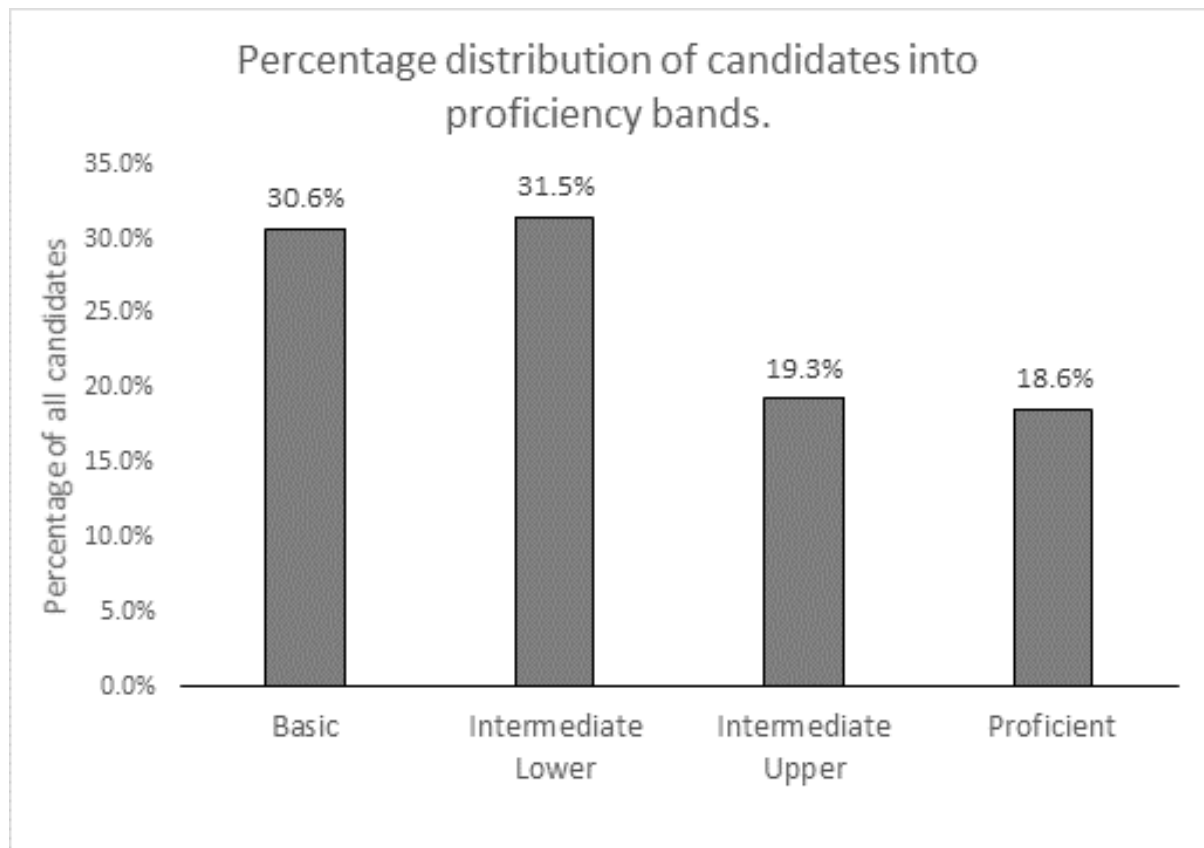


Figure 1: Percentage of total sample in each proficiency band

Figure 2 shows the distribution of scores for all the candidates and in each proficiency band. The median score for the whole sample was well within the Intermediate Lower band, while the upper quartile was within the Intermediate Upper band. While the scores in the two Intermediate bands were fairly evenly spread out across the range of values for those bands, the Basic scores were concentrated towards the upper end of the Basic band and the Proficient scores were concentrated towards the lower end of the Proficient band.

The differences between the performances of male and female candidates are shown in Figures 3 and 4. From Figure 3 we see that a greater proportion of females than males were placed in the lower two proficiency bands, with a correspondingly lower proportion of females than males in the Proficient band. The differences between the distributions of overall scores for the whole test are shown in Figure 4 where we see that in general the scores for females

were lower, with the median for females being well within the Intermediate Lower band, while the median value of males is just below the top of this band.

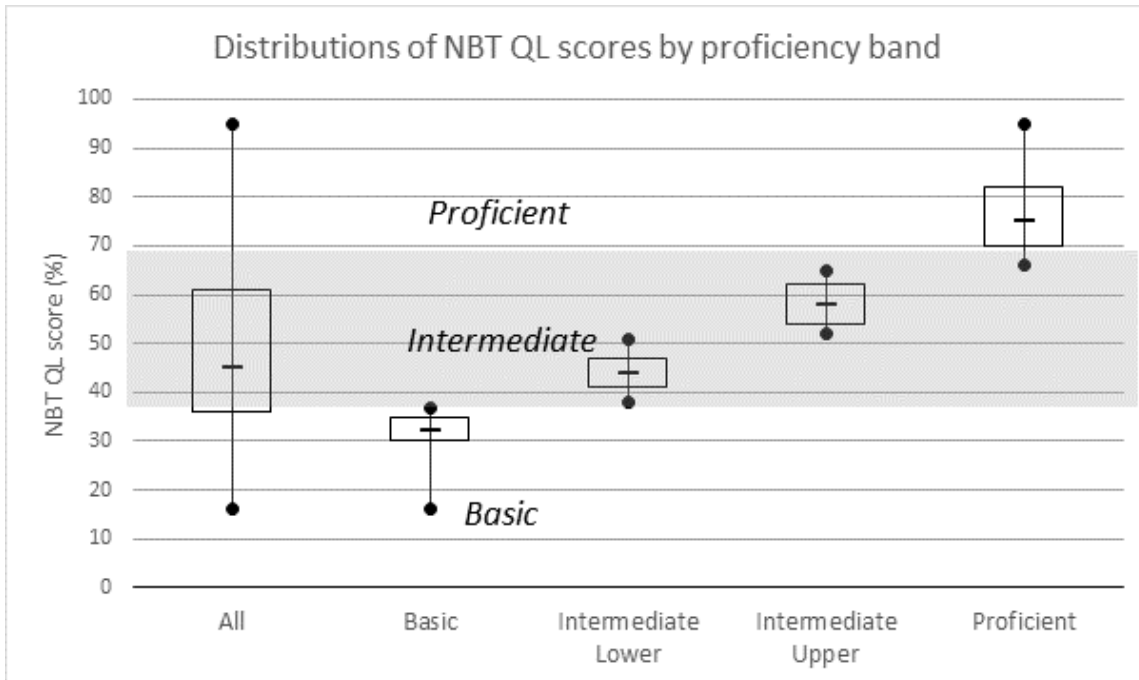


Figure 2: Distribution of scores for each proficiency band

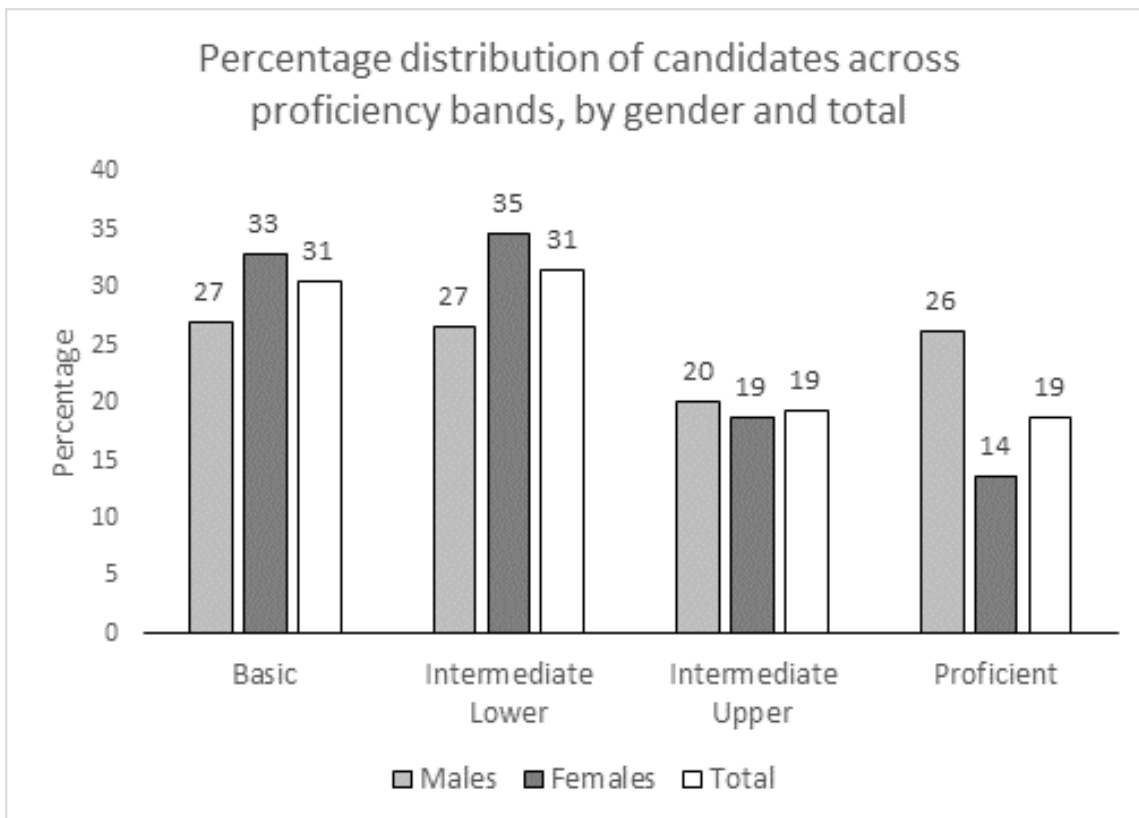


Figure 3: Percentage of males and percentage of females and percentage of whole sample in each proficiency band.

Distributions of scores for subgroups of test items.

Based on a judgement of the main ideas addressed, each item was assigned to one of the ‘mathematical and statistical ideas’ categories defined in the test construct, with the exception of six items which were classified as addressing two of these ideas in equal measure, namely data and quantity (see the description of the categories in the section under the heading ‘NBTP QL test item subgroup analysis’ above). This results in six disjoint subgroups of items, for which the distribution of the performances of the entire sample are shown in Figure 5. There was also a subgroup of items that addressed the idea of chance and uncertainty, but there were too few of these to yield meaningful results and they are therefore not reflected in the chart.

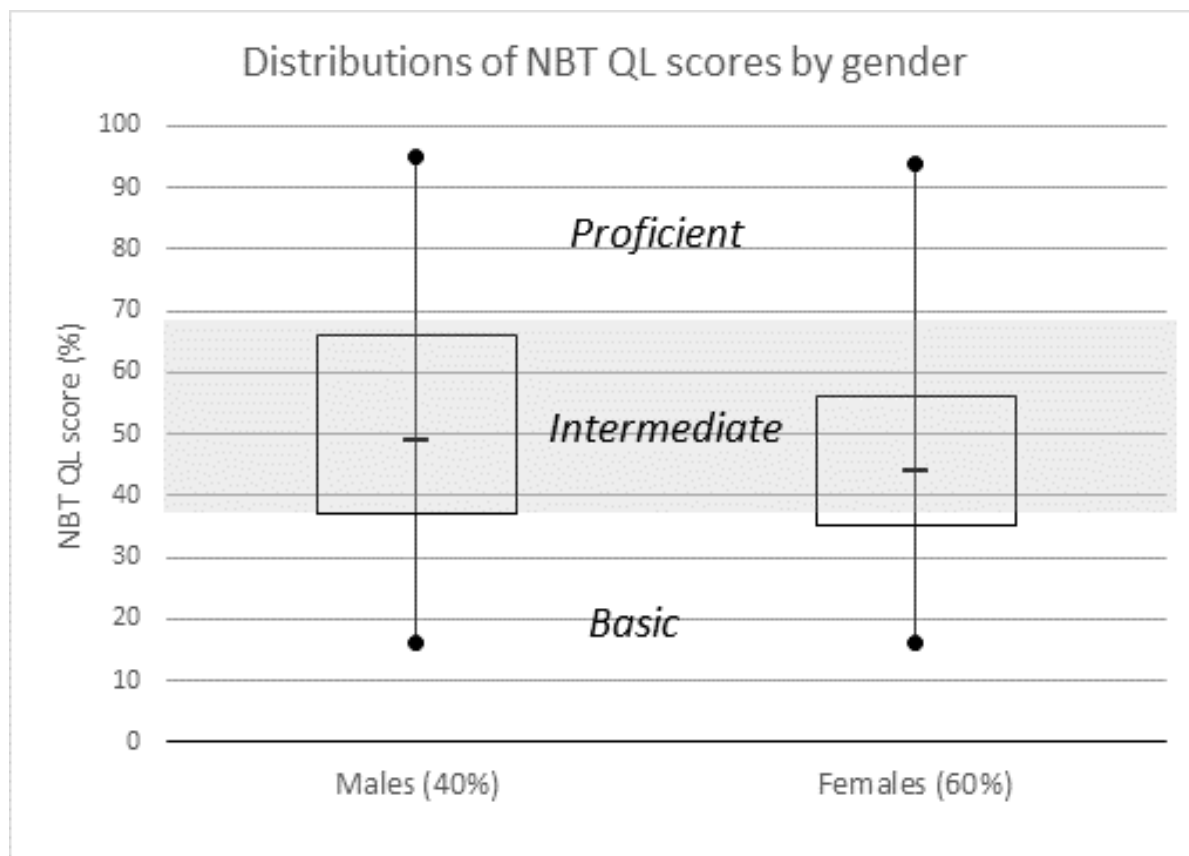


Figure 4: Distribution of scores for males and for females.

The distributions for the subgroups showed some variation but their medians were fairly similar to each other and to the overall distribution (see Figure 2). Three quarters of the candidates scored below the Proficient level in all subgroups except the subgroup of items that mainly addresses the idea of quantity. It is striking that those items which involve both the ideas of data and quantity appear to be most challenging, with three quarters of the candidates scoring less than 50 per cent on this subgroup. This probably reflects the relatively high difficulty level of items that require some combination of selecting appropriate data from a chart or table with doing a calculation or estimation. The subgroup on which candidates performed best consists

of those items that were judged to address mainly the idea of quantity, where three quarters of the candidates scored 44 per cent or more. Once again this could be a reflection of the difficulty level of the items, since this subgroup is most likely to contain some of the easiest single-step items. Candidates also performed relatively poorly on the items that mainly address the idea of relationships, which is not unexpected as the items in this subgroup tend to be somewhat more mathematical and may involve the use of simple mathematical symbols.

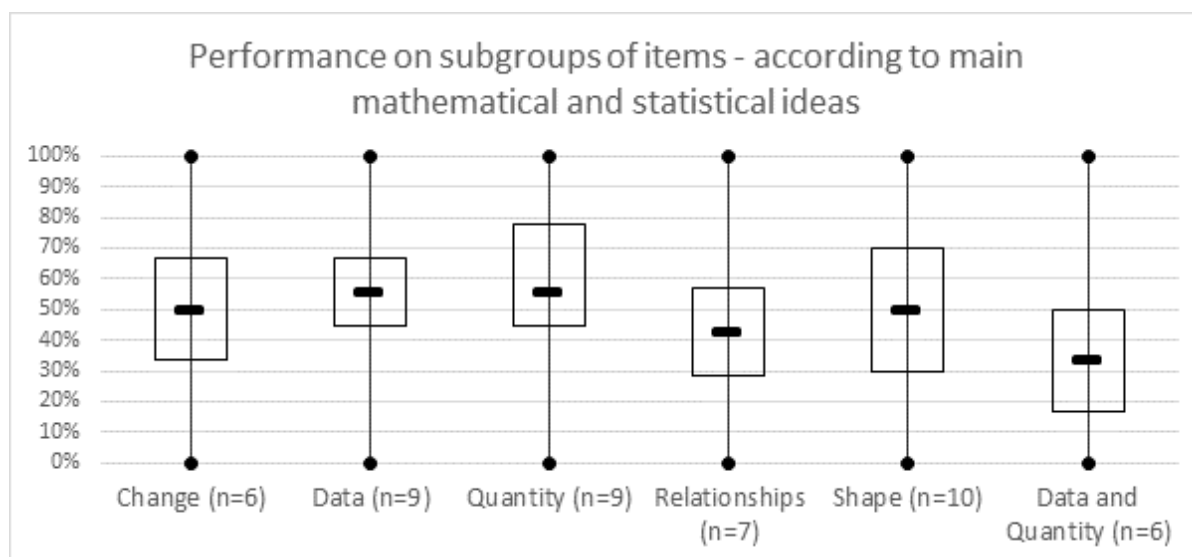


Figure 5: Distribution of scores for entire sample on subgroups of the test items defined in terms of the main mathematical and statistical ideas addressed by the items.

We have seen that interpreting the performance of the candidates on subgroups of items defined by their main mathematical and statistical ideas was subject to the confounding effect of the cognitive demand of items, which this classification did not take into account. For this reason, the analysis based on classifying the items into subgroups on the basis of the main competency the items were designed to assess is also useful (see the description of the categories in the section under the heading 'NBTP QL test item subgroup analysis' above). The distributions of scores for these subgroups are shown in Figure 6.

In all subgroups except 'using representations', the upper quartile was well below the top of the Intermediate Upper band, indicating that well over three quarters of the sub-scores were not in the Proficient band. The subgroup on which the performance was the weakest is 'computing', with a median of 33 per cent and 25 per cent of candidates scoring less than 17 per cent. The strongest performance was on the subgroup 'deriving information from data representations'. This is the only subgroup for which the upper quartile is above the top of the Intermediate band. These questions tend to be more direct and require less interpretation, and also presumably draw on skills that are better developed at school. The median scores for

‘reasoning’, ‘translating’ and ‘extrapolating and (higher order thinking competencies) were all similar and were just above the bottom of the Intermediate Lower band.

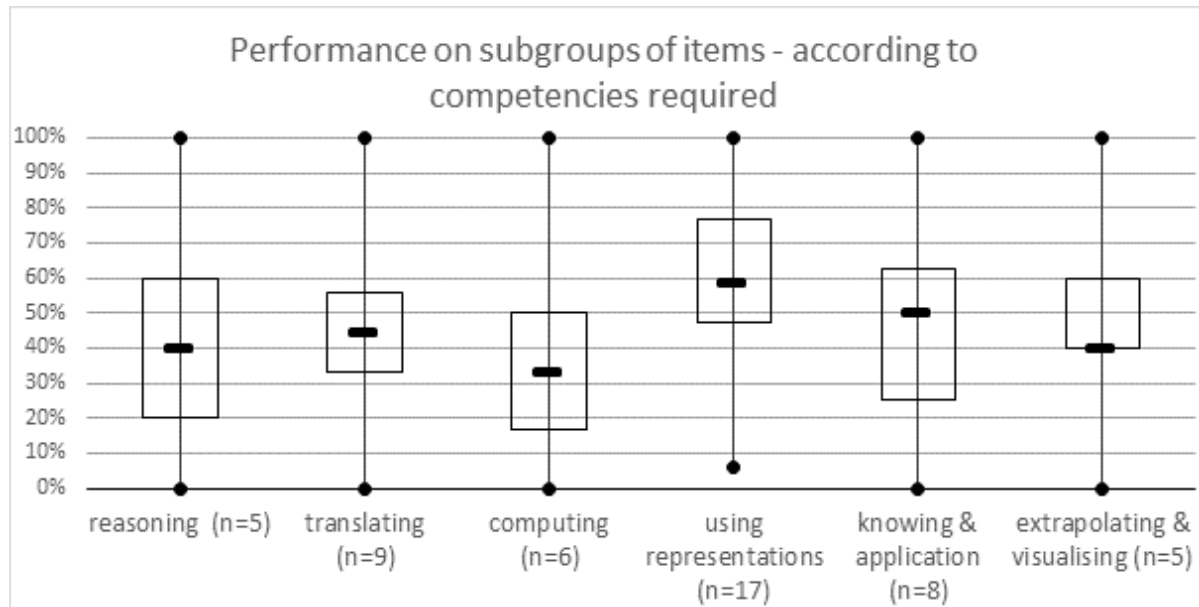


Figure 6: Distribution of scores for entire sample on subgroups of the test items defined in terms of the competencies required by the items. (Category labels are abbreviated as follows: reasoning – ‘reasoning and synthesising’; translating – ‘translating between different representations’; computing – ‘interpreting problem statements and computing’; using representations – ‘deriving information from data representations’; knowing & application – ‘knowing simple facts and applying them’)

In this section we present some examples to illustrate how insights into students’ abilities can be obtained by a close examination of the proportions of candidates who choose different alternative answers on individual items. Alternative answers to the correct one often encapsulate misconceptions or erroneous thinking. The difficulty in reporting this kind of information is that for security reasons one cannot widely publicise the actual test items, but understanding student performance at this level depends on knowing fairly precisely the structure and content of the items. So we will attempt to illustrate the diagnostic use of analysis of candidates’ choices of alternatives without showing actual items. We will however use contexts from higher education learning materials for this purpose. The discussion will focus on the abilities of the candidates in the Intermediate categories because this is where the majority of candidates, and hence of students entering higher education, is located.

Example 1: Interpreting percentage values in a table.

Table 4 comes from a reading provided for 1st year medical students. Students need to understand that, for example, the percentage at the top of the second column means: ‘76.9 per cent of the homicides of people under 15 years in age were males’, not ‘76.9 per cent of the male homicides were under 15’ or ‘76.9 per cent of under 15 year old males were homicides’

or some other formulation. This understanding depends on identifying which percentages in the table add up to 100 per cent and hence what the denominator is in the percentage calculation.

Figure 7 shows the results for an item in the NBTP QL test referring to a table with similar data as in Table 4. The chart shows the choices made by candidates in the different proficiency bands separately (based on their performance on the whole QL test) and the choices for the whole group at the top.

Table 4: Example of data table containing percentages that require interpretation

Reproduction of Table 1 from (Peden and Butchart 1999, 335)				
Homicide and suicide by gender and age (preliminary NMSS data, first quarter 1999)				
Age ranges	Homicides		Suicides	
	Males	Females	Males	Females
< 15	10 (76.9%)	3 (23.1%)	1 (50.0%)	1 (50.0%)
15 - 24	166 (91.2%)	16 (8.8%)	22 (91.7%)	2 (8.3%)
25 - 34	231 (91.3%)	22 (8.7%)	23 (85.2%)	4 (14.8%)
35 - 44	141 (82.9%)	29 (17.1%)	11 (68.8%)	5 (31.2%)
45 - 54	48 (84.2%)	9 (15.8%)	9 (81.8%)	2 (18.2%)
55 - 64	20 (87.0%)	3 (13.0%)	8 (100%)	0
65 +	3 (42.9%)	4 (17.1%)	5 (62.5%)	3 (37.5%)

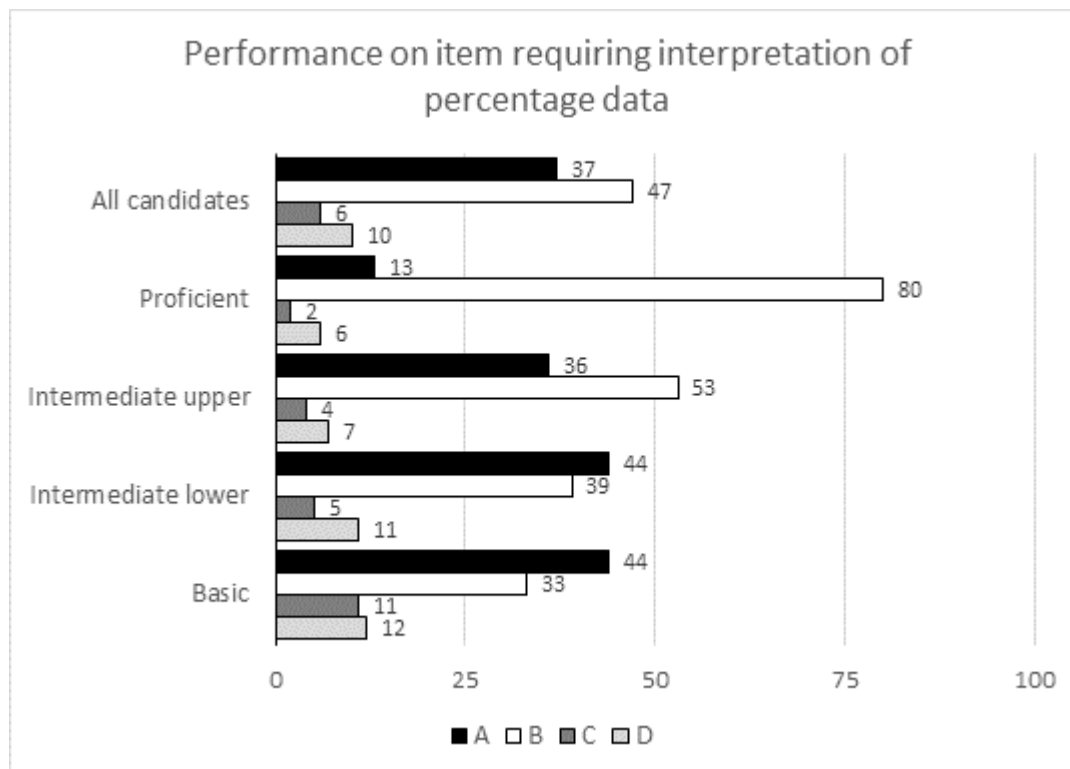


Figure 7: Pattern of performance on an item requiring interpretation of percentage data from a table like the one in Table 4 above. For that example alternative answers A, B, C and D would be equivalent to the following: A - 76.9% of the male homicides were under 15 years of age; B - 76.9% of the homicides of under 15 year olds were males; C - 76.9% of the homicides were of males under 15 years of age; D - 76.9% of the homicides were under 15 years of age.

The part of the chart for 'All candidates' shows that less than half could identify the correct description of the meaning of a particular percentage value in a table (alternative B) and only about half in the Intermediate Upper category and about 40 per cent in the Intermediate Lower category could do this.

Alternative A was very popular with the weaker candidates (for example 44% and 36% of the Intermediate Lower and Upper categories respectively chose this option). In the above example, this would be equivalent to their choosing '76.9 per cent of the male homicides were under 15 years of age' which involves the correct row and column categories, but does not recognise the correct denominator for the percentage calculation. These candidates were unable to recognise how to tell from the table where the percentages add up to 100 per cent and deduce from this what the 'whole' is that the percentages are a 'part' of.

The pattern of performance on this task indicates that many students will have difficulties with quantitative language and in particular the language used to describe percentages. It also indicates that there are general difficulties with interpreting tabulated data, especially when it involves percentages.

Example 2: Definition of percentage change

This example shows that it would be problematic to assume that students understand what it means to say that a quantity has experienced a particular percentage change, although percentage change is a concept that is very commonly encountered in academic materials. Figure 8 shows that only 35 per cent of all candidates recognised that percentage change is calculated as a percentage of the initial value (alternative B), while a third of all candidates (and 28 per cent and 42 per cent in the Intermediate Upper and Lower bands respectively) selected an answer based on the change calculated as a percentage of the final value (alternative A).

Example 3: Radius of circular object drawn to scale.

Figure 9 shows the pattern of responses for a question that involved using a scale diagram to find the radius of a round shape. Less than half of all the candidates (and 39% and 60% of Intermediate Lower and Upper respectively) chose the correct answer (alternative C), while across all proficiency bands about one third chose the answer that gave the diameter (alternative B). Either many candidates do not know the definition of radius and diameter, or there was a very great tendency not to pay attention to precision in answering questions (regardless of proficiency).

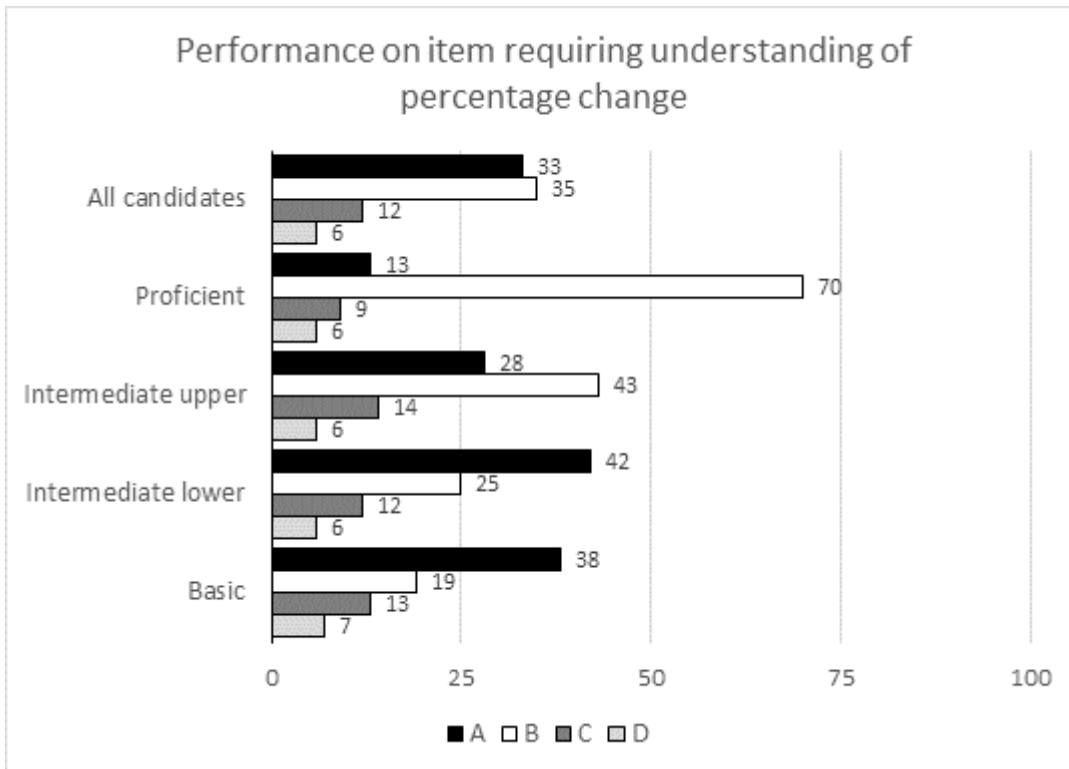


Figure 8: Pattern of performance on an item requiring understanding of the definition of percentage change

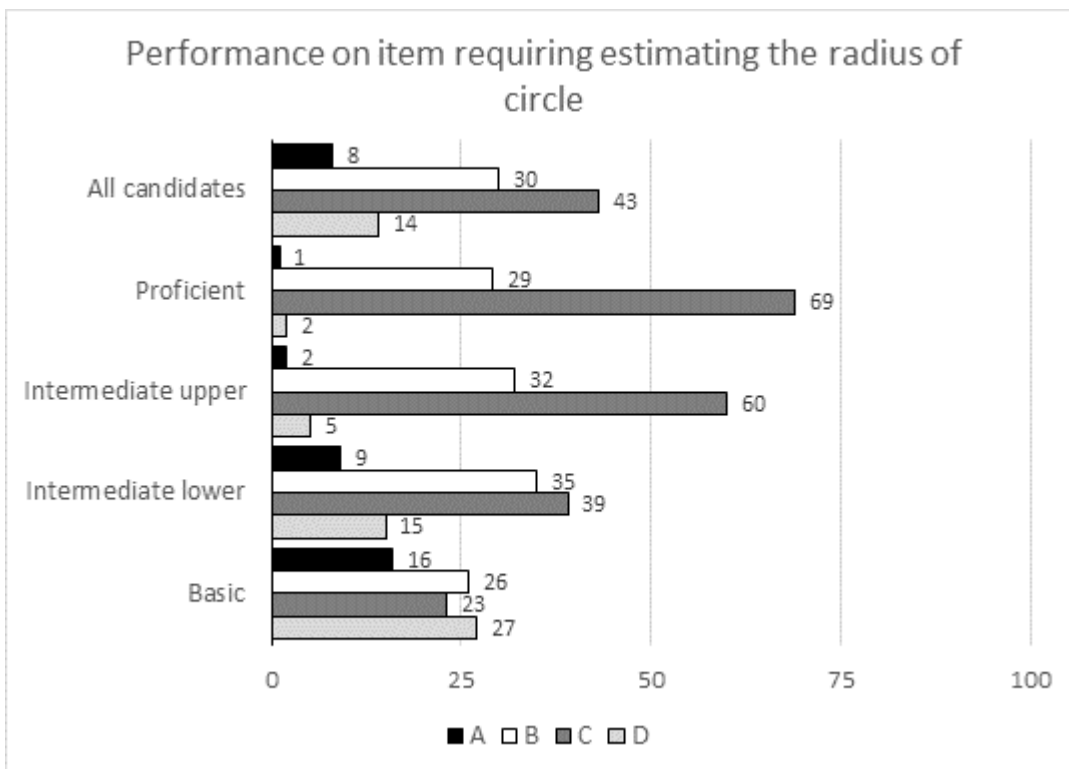


Figure 9: Pattern of performance on an item requiring estimating the radius of a circular shape from a scale diagram

Example 4: Proportional reasoning

The following example illustrates the ‘blindness’ that many students have to the distinction between absolute and relative quantities, the language used to make this distinction and the ability to reason with proportions. Given data for the total population of a few countries (for example in a table) and the proportion of the population that is under 15 years of age (for example in a stacked bar chart), a question might ask which country has the greatest number of people under 15 year of age. In a situation where the proportions are similar, but one of the countries has a vastly bigger population than the others, it is possible to reason that this country must also have the largest number of people under 15 years of age.

Figure 10 represents the results for a task similar to this one, where alternative A represents the correct answer and alternative C represents the greatest proportion (not number). Only 19 per cent of all candidates succeeded in choosing the greatest number and only half of the proficient candidates could do so. By far the greatest number of candidates (well over 80% in the Intermediate categories) chose the largest proportion (alternative C).

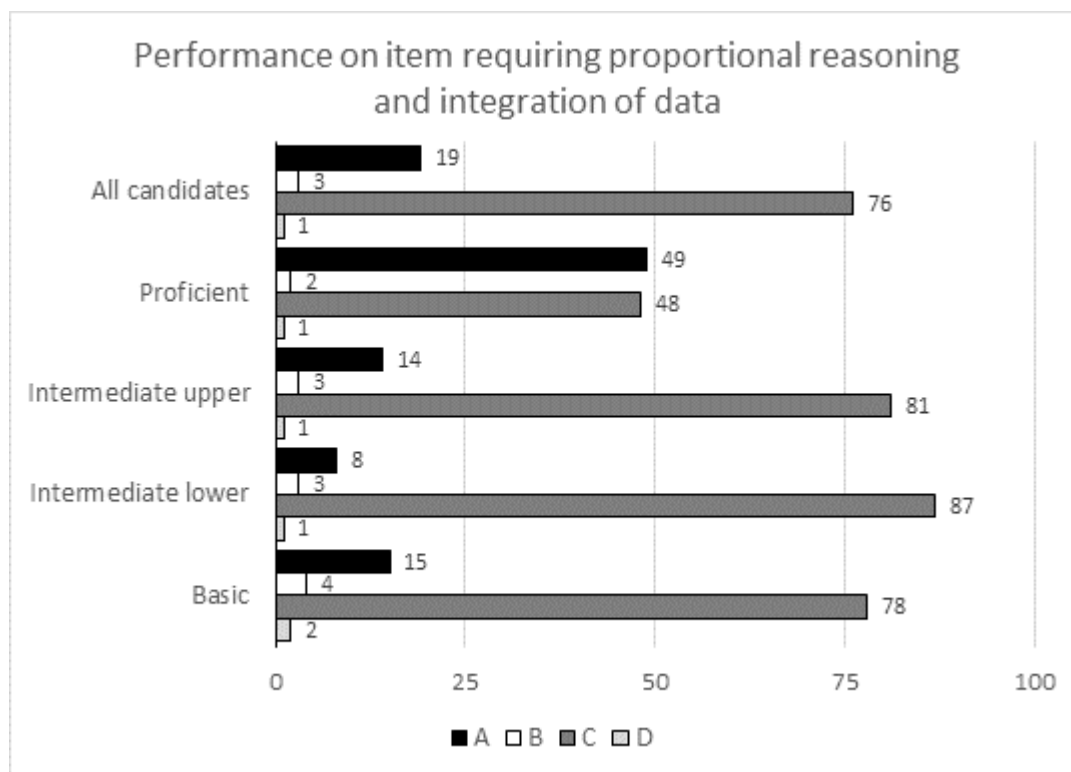


Figure 10: Pattern of performance on an item requiring integration of data from two sources and proportional reasoning.

The pattern of performance on this task indicates a lack of understanding of the language used to describe absolute and relative quantities and a lack of proportional reasoning ability. In

addition, it also seems that candidates did not recognise the need to integrate information from more than one data source.

DISCUSSION AND CONCLUSION

In this article the results are presented for a large sample ($n = 6\,326$) of school-leavers from across South Africa who wrote one version of the NBTP QL test in 2011 for entry into higher education in 2012. Over 60 per cent of the scores for the test were in the lower two proficiency bands, with 30 per cent in the Basic band and less than 20 per cent in the Proficient band. This suggests that the majority of candidates aiming to enter higher education are in need of some kind of supplementary support for developing their quantitative literacy, while at least 30 per cent would require extensive support. In general, the overall scores were lower for females than for males, with the median for females being well within the Intermediate Lower band while the median value for males is just below the top of this band. The sample of candidates considered here contained a disproportionately high fraction of females (60%), which is consistent with national cohorts.

The scores on subgroups of items classified according to the main mathematical and statistical ideas they address are also considered. Three quarters of the candidates scored below the Proficient level in all of these subgroups except the one that mainly addressed the idea of quantity. The weakest performance was in the subgroup that required candidates to integrate, by obtaining information from a data source to use in a calculation. Integration is a fairly high order thinking skill, but one that plays a very important role in quantitative literacy practice in academic disciplines.

The distributions of scores on subgroups of items classified according to the main competencies they were designed to assess are also presented. In all subgroups except 'deriving information from data representations', well over three quarters of these sub-scores were not in the Proficient band. The subgroup on which the performance was the weakest was 'computing', which we suggest is a result of candidates' dependence on calculators and consequent poor number sense. The strongest performance was on the subgroup 'deriving information from data representations'. This kind of analysis of the scores on subgroups of items is useful for providing diagnostic information that allows for interventions to be designed that focus on the specific needs of students in terms of both mathematical content and competencies to be developed.

In the final section of the article, some examples of diagnostic information that can be derived from close examination of patterns of performance on individual items are also

presented. These indicate that many candidates had difficulty with quantitative language and in particular the language used to describe percentages and absolute and relative quantities. Other insights revealed by these examples are that candidates generally had difficulties with interpreting tabulated data, especially when it involved percentages, that many lacked proportional reasoning ability, and that many did not understand the concept of percentage change. These are all fundamental competencies integral to quantitative literacy practice in most academic disciplines. They are also essential for critical citizenship in a society where many arguments and understandings depend heavily on quantitative data. It would therefore be important to foster these competencies as graduate attributes in higher education.

These results stress that for the majority of students in higher education there is a need for their quantitative literacy to be explicitly developed. Each academic discipline will make different demands in terms of quantitative literacy (and other academic literacies). The NBTP test results can be presented in a disaggregated way for students on different programmes and in different courses. There is thus significant potential for the NBTP to inform appropriate curriculum development across the South African HE sector. This is in fact the focus of the Standard Bank Teaching and Learning (T&L) project which supports the National Benchmark Tests Project in meeting its objective to provide diagnostic information to improve curriculum development. The T&L project was initiated in 2014 and aims to facilitate the use of NBTP data for improving teaching and learning at course, programme, faculty and institutional levels for all HE institutions in SA.

REFERENCES

- Boughey, C. 2009. Understanding teaching and learning at foundation level: A 'critical' imperative? In *Beyond the university gates: Provision of extended curriculum programmes in South Africa. Proceedings of the January 2009 Rhodes University Foundation Seminar*, ed. C. Hutchins and J. Garraway, 4–7. Grahamstown: Rhodes University. http://www.cput.ac.za/storage/services/fundani/beyond_the_university_gates.pdf#page=9 (accessed 24 February 2015).
- Centre for Educational Testing for Access and Placement 2015. *National Benchmark Tests Project National Report: 2015 Intake Cycle*. Cape Town: Centre for Educational Testing for Access and Placement. <http://tinyurl.com/onhjgdk> (accessed 24 February 2015).
- Chapman, A. and A. Lee. 1990. Rethinking literacy and numeracy. *Australian Journal of Education* 34(3): 277–289. <http://dx.doi.org/10.1177/000494419003400305> (accessed 2 June 2015).
- Council on Higher Education. 2013. *A Proposal for Undergraduate Curriculum Reform in South Africa: The Case for a Flexible Curriculum Structure*. Pretoria: Council on Higher Education. http://www.che.ac.za/sites/default/files/publications/Full_Report.pdf (accessed 2 June 2015).
- Department of Education, South Africa. 2003. *National Curriculum Statement Grades 10–12 (General). Mathematical Literacy*. Pretoria: Department of Education.
- Foxcroft, C. 2006. The nature of benchmark tests. In *Access and entry level benchmarks, the National Benchmark Tests Project*, ed. H. Griesel, 7–16. Pretoria: Higher Education South Africa. http://www.hesa.org.za/sites/hesa.org.za/files/2006_HESA_Access_and_Entry_Level_Benchmarks

- .pdf (accessed 24 February, 2015).
- Frith, V. and R. Prince. 2006. Quantitative literacy. In *Access and entry level benchmarks, the National Benchmark Tests Project*, ed. H. Griesel, 28–34; 47–54. Pretoria: Higher Education South Africa. [http://www.hesa.org.za/sites/hesa.org.za/files/2006_HESA_Access and Entry Level Benchmarks.pdf](http://www.hesa.org.za/sites/hesa.org.za/files/2006_HESA_Access_and_Entry_Level_Benchmarks.pdf) (accessed 24 February 2015).
- Gal, I., M. van Groenestijn, M. Manly, M. J. Schmitt and D. Tout. 2005. Adult numeracy and its assessment in the ALL Survey: A conceptual framework and pilot results. In *International Adult Literacy Survey. Measuring adult literacy and life skills: New frameworks for assessment*, ed. by T. Scott Murray, Y. Clermont and M. Binkley, 137–191. Ottawa: Statistics Canada. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.4652&rep=rep1&type=pdf> (accessed 24 February 2015).
- Griesel, H. 2006. *Access and Entry Level Benchmarks, the National Benchmark Tests Project*. Pretoria: Higher Education South Africa. [http://www.hesa.org.za/sites/hesa.org.za/files/2006_HESA_Access and Entry Level Benchmarks.pdf](http://www.hesa.org.za/sites/hesa.org.za/files/2006_HESA_Access_and_Entry_Level_Benchmarks.pdf) (accessed 24 February 2015).
- Hambleton, R. K. and M. J. Pitoniak. 2006. Setting performance standards. In *Educational measurement*, ed. R. L. Brennan, 433–470. 4th ed. Westport, CT: Greenwood/Praeger.
- Holland, P. W. and N. J. Dorans. 2006. Linking and equating. In *Educational measurement*, ed. R. L. Brennan, 187–220. 4th ed. Westport, CT: Greenwood/Praeger.
- Hughes-Halett, D. 2001. Achieving numeracy: The challenge of implementation. In *Mathematics and democracy: The case for quantitative literacy*, ed. L. A. Steen, 93–98. USA: The National Council on Education and the Disciplines.
- Johnston, B. 2002. *Numeracy in the making: Twenty years of Australian adult numeracy. An investigation by the New South Wales Centre, Adult Literacy and Numeracy Australian Research Consortium*. Sydney: University of Technology, Sydney.
- Johnston, B. 2007. Critical numeracy? In *The adult numeracy handbook. Reframing adult numeracy in Australia*, ed. S. Kelly, B. Johnston and K. Yasukawa, 50–56. Sydney, NSW: Adult Literacy and Numeracy Australian Research Consortium.
- Kelly, S., B. Johnston and M. Baynham. 2007. The concept of numeracy as social practice. In *The adult numeracy handbook. Reframing adult numeracy in Australia*, ed. S. Kelly, B. Johnston and K. Yasukawa, 35–49. Sydney, NSW: Adult Literacy and Numeracy Australian Research Consortium.
- McKenna, S. 2009. Cracking the code of academic literacy: An ideological task. In *Beyond the university gates: Provision of extended curriculum programmes in South Africa. Proceedings of the January 2009 Rhodes University Foundation Seminar*, ed. C. Hutchins and J. Garraway, 8–15. Grahamstown: Rhodes University. http://www.cput.ac.za/storage/services/fundani/beyond_the_university_gates.pdf#page=9 (accessed 24 February 2015).
- Peden, M. and A. Butchart. 1999. Trauma and injury. In *South African Health Review 1999*, ed. N. Crisp and A. Ntuli, 331–344. Durban: Health Systems Trust. http://www.hst.org.za/sites/default/files/chapter24_99.pdf (accessed 23 June 2015).
- Scott, I., N. Yeld and J. Hendry. 2007. *Higher Education Monitor No. 6: A Case for Improving Teaching and Learning in South African Higher Education*. Pretoria: The Council on Higher Education. http://www.che.ac.za/sites/default/files/publications/HE_Monitor_6_ITLS_Oct2007_0.pdf (accessed 24 February 2015).
- Steen, L. A. 2004. *Achieving quantitative literacy: An urgent challenge for higher education*. Washington D.C.: The Mathematical Association of America.
- Street, B. and D. Baker. 2006. So, what about multimodal numeracies? In *Travel Notes from the New Literacy Studies*, ed. K. Pahl and J. Rowsell, 219–233. Clevedon: Multilingual Matters Ltd.
- Street, B. 2005. Applying new literacy studies to numeracy as social practice. In *Urban literacy. Communication, identity and learning in development contexts*, ed. A. Rogers, 87–96. Hamburg: UNESCO Institute for Education.
- Yeld, N. 2009. The National Benchmark Tests Project: Addressing student educational needs in the

tertiary education system. In *Recession and recovery*, ed. J. Hofmeyer, 76–83. Cape Town: Institute for Justice and Reconciliation. http://www.ijr.org.za/publications/pdfs/IJR_TA_Chapter3.pdf (accessed 24 February 2015).

Yen, W. M. and A. R. Fitzpatrick. 2006. Item response theory. In *Educational measurement*, ed. R. L. Brennan, 111–153. 4th ed. Westport, CT: Greenwood/Praeger.