

THE NATIONAL BASIC SKILLS SURVEY OF ADULTS IN ENGLAND 2002-3: THE NUMERACY SURVEY TO DATE

John Gillespie

<John.Gillespie(at)nottingham.ac.uk>

@ is replaced by (at) to stop the automatic garnering of email addresses by Spam factories - Editor

Centre for Developing and Evaluating Lifelong Learning,
School of Education, University of Nottingham (CDELL)

Introduction

The Adult Basic Skills Strategy Unit in the Department for Education and Skills commissioned CDELL to plan and devise two collections of assessment items to form the basis of the National Survey of literacy and numeracy. A profiled sample of 10000 adults aged 16 – 64 are being interviewed in their homes using laptops to present the survey items and store individuals' responses. The aim is to produce a national profile of adult literacy and numeracy competence over five Literacy and Numeracy Core Curriculum for England and Wales. It is hoped that the results of the surveyhis paper concerns the numeracy part of the survey. The author led the CDELL numeracy team.

Partner organisations

The survey is being carried out by the British Market Research Bureau (BMRB). Associated software has been designed by Bradford Technology Limited (BTL). The numeracy and literacy items and the design of the two surveys are the responsibility of two teams from CDELL.

The main aim of the survey

Much learning and teaching of numeracy with adults in England is now based on the Adult Numeracy Core Curriculum (Basic Skills Agency, 2001). This curriculum is presented in five levels – from lowest to highest these are Entry Levels 1, 2 and 3 and then Levels 1 and 2. Level 2 contains content corresponding to the Key Skills Application of Number level 2 specifications (t is broadly comparable in technical demand to aspects of Intermediate level GCSE mathematics, while Entry Level 1 contains content comparable to the curricula and attainment of many six and seven year olds.

A main aim of the survey is to produce national estimates, for the first time, of the proportions of the adult population of England currently at each of these levels, which could then be presented by age, sex, location and socio-economic grouping, so as to act as evidence for future comparisons and to inform future educational and training planning and interventions aimed at raising literacy and numeracy levels in England.

Ongoing nature of the work

The survey of a representative sample of approximately 10,000 adults in England commenced in July 2002. It is not likely to be concluded and the results presented by the Department for Education and Skills until the summer of 2003. Until then, individual items may not be made public, nor may any emerging results be published. However, aspects of the survey design, the considerations that led to them and other details of the survey process are not restricted; these form the subject of this paper.

Considerations of the survey population

- In devising the numeracy survey, the team had several considerations in mind
- The full range of numeracy ability was likely to be encountered - ranging from individuals who might have difficulty even with working with two digit numbers, up to those who were of degree standard in mathematics.
- Respondents would have no reason to co-operate with the survey apart from an altruistic one, personal interest and a modest complimentary payment on completion.
- For many respondents, their previous experience of working with mathematics might well have been unpleasant, making them reluctant to take part and likely to be easily discouraged by questions that were too demanding.
- In addition, experience from adult numeracy students suggested that many respondents would show a 'spiky profile' of competence – perhaps, for example. quite comfortable with arithmetic of money, but having difficulties with percentages, in interpreting graphs to respond to such profiles.
- Finally, the time available to undertake the numeracy part of the survey was to be no more than 30 minutes. This meant that in fixing the extent of the survey, the team had in mind an average total time for the numeracy section of no more than 20 minutes.

Initial design considerations and features

The survey was to be carried out using multiple-choice items presented to respondents by lap-top computers.

The project team commenced work in December 2001, so that time to carry out item design, piloting and other research was very constrained. A specification for the numeracy items was drawn up in January 2002, with expert advice from Dr Diana Coben , Dr Jeff Evans, Professor Margaret Brown, Dr Alison Tomlin and others.

The items were designed by a team of three writers - all experienced in adult numeracy assessment. A proportion of the items at the upper two levels (Levels 1 and 2) were required to be closely based on items previously used in Adult Numeracy assessments, adapted to fit the survey requirements and screen layout. All items for the lower three levels (Entry levels 1, 2 and 3) were new. In designing the items, the authors took account of items used in other numeracy surveys of standing, including DfEE (1999), Elkinsmyth and Bynner (1994), IALS test items, van den Heuven-Panhuizen, M (1994 and 1996) and PISA - Programme for International Student Assessment (2001). In addition, ideas and approaches outlined in recent research into aspects of adult numeracy were referred to (Coben et al. 2000).

Piloting took place with groups of adult numeracy students and their tutors, enabling improvements to the wording and presentation of items to be carried out. Each item was then re-checked against the Core Curriculum statements for levels above and below the intended level of the item to ensure that the item best fitted its intended level.

Several innovative features have been included which the project team feel have contributed to the emerging success of the survey process. In particular, a series of algorithms was developed by the author to route the individual respondent to items at an appropriate level for that person, based on their previous responses, in the style of adaptive testing.

Respondents are presented with items in seven groups or 'steps'. Each of these seven steps targets different aspects of numeracy. In the first step, all respondents meet the same four items, two at Entry Level 1 and one each at Entry Levels 2 and 3. These were deliberately chosen so as to present familiar and straightforward tasks to all respondents. Based on their performance, respondents are then directed to one of three overlapping groups of five items, forming Step 2, with items ranging from Entry Level 1 to Level 2. Depending on their performance on these, the algorithm takes respondents to two items of an appropriate level in Step 3; these range from two at Entry Level 1 to two at the top level - Level 2. Again depending on their performance on these, the algorithm takes respondents to two appropriate items in Step 4. This is repeated up to Step 7 so that each respondent encounters 19 items in all, from a total of 48 items altogether.

Table 1 lists the 48 items analysed by general topic, step and level. An extract from the progression algorithm is shown in Figure A1 (see end of article). The numbers in boxes represent the items numbers and the arrows show progression routes depending on correct (C) and not correct (N) responses at each Step. The algorithm patterns for Steps 5, 6 and 7 are similar to those for Step 4.

Table 1. Analysis of Levels and Steps of items.

Level	E1	E2	E3	L1	L2	of which the respondent is presented with
Step number and topic(s)						
Step 1 Basic money calculations	11 12	13	14			4
Step 2 Whole number calculations and time	21 22	23 34	25 26	27 28	29	5
Step 3 Measures and proportion	31 32	33	34	35	36 37	2
Step 4 Weight and scales	41 42	43	44	45	46 47	2
Step 5 Length and scaling	51 52	53	54	55	56 57	2
Step 6 Charts and data	61 62	63	64	65	66 67	2
Step 7 Money calculations	71 72	73	74	75	76 77	2
Total number of items at this level	14	8	8	7	11	
Total number of items					48	19

Owing to cost considerations and requirements from DfES, not all the advisers' recommendations could be acted upon. Those not incorporated included that the items should have a voice-over option to assist with reading problems, and that calculators should be permitted for some items. The former meant that the text and layout on the screen had to be as straightforward and easy to read as possible. There was concern that the reading requirement would exclude a small but significant group of potential respondents, but subsequent experience appears to indicate that very few respondents were actually so excluded.

The conduct of the survey

In the survey, the numeracy items are presented to respondents by trained BMRB interviewers. The interviewer typically sits alongside the respondent so that they can both see the lap-top screen. Before the first survey item is shown, two pre-survey items are presented to respondents to show them the styles of item they will be meeting and to enable the interviewer to explain what will be happening. The first survey item is then shown. The respondent reads the item, then selects from typically four alternative answers. The interviewer then inputs this choice into the lap-top: the next item is then selected automatically according to the algorithm and displayed. The interviewer's role is to input the respondents' choices correctly: the interviewer may not read out a question or provide hints of any sort.

The sequence of items shown and the respondents' response choices and times are recorded automatically.

Opportunities presented by the use of lap-top computers

Personal observation confirmed the reports from interviewers from the first batch of surveys that respondents reacted well to the use of lap-tops. Typically, the lap-top was seen as a neutral question-setter with the interviewer being viewed as 'on the same side' as the respondee of working from it from previous learning experiences.

Crucially, the adaptive design has meant that respondents are presented with items by and large appropriate to their levels of numeracy ability, while also reacting to individual respondents' areas of facility or difficulty.

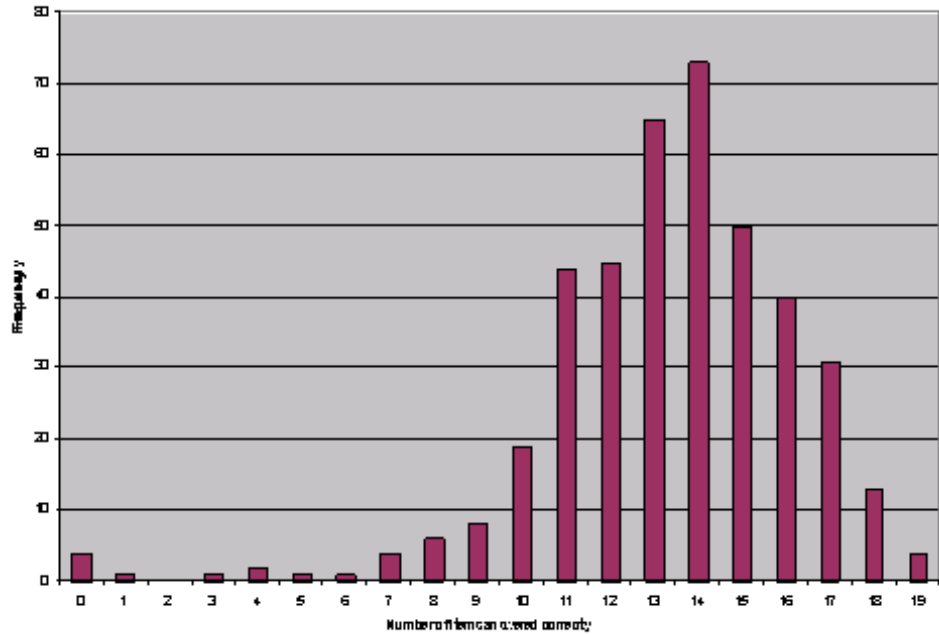
Success rates for individuals - drop-out rates

A main design objective was to base the estimates of level on what respondents could do, rather than what they couldn't. A subsidiary objective was to encourage and motivate respondents through their positive reactions to the survey early on. It has been expected to have regarding exposing their numeracy skills. Indications to date are that both objectives have been substantially satisfied.

Figure 1 shows the frequencies of different numbers of correct answers from the first 412 respondents. The mean number of correct responses was 13.3 while over 90% of all respondents selected ten or more correct answers from the possible 19.

In addition, to date less than 2% of respondents have failed to address the full set of 19 items they were presented with. We believe that the adaptive nature of the survey was a major contributor to these gratifying results.

Figure 1 Distribution of number of items answered correctly



Scoring and assessing

The survey is designed to estimate the proportion of respondents at each of five levels:

- at or below Entry Level 1
- at Entry Level 2
- at Entry Level 3
- at Level 1
- at or above Level 2

Individuals would be likely to have performed at different levels of competence on different topic areas. Thus, many respondents' performance records would show a series of correct responses to a set of items set at different levels. What would be the most appropriate method to converting individual performance records into estimates of overall level? Would overall level be best measured by the level of the final two items successfully tackled - that is by the level of successful performance at Step 7? Or should it be based on mean or median performance on the final ten items from Steps 3 to 7? Or should it be based on summing overall performance (scoring 1 for a correct EL1 response up to 5 for a correct L2 response)? In all, five alternative schemes for setting overall level were trialled and these were compared against detailed analysis of individuals' performances from the first 189 respondents' results

The method finally chosen was to sum overall performance, as this took into account all aspects of the respondent's performance. This led to the setting of threshold scores for minimum scores to achieve a particular level. These thresholds were carefully chosen after scrutiny of individual performances from the first 412 respondents and of the performance of individual items. Thus the few items which turned out to have very low or very high facility levels could be allowed for, while final decisions on thresholds for the five levels of performance could be deferred until after the data collection had been completed.

To confirm this method of estimating level, levels were then re-calculated using level estimates based on performance on the 'final ten' and 'final eight' items. Overall proportions using the three methods were found to be very close to each other.

Likely publication of results

It is unlikely that the data collection will be completed before late spring 2003. It is therefore most unlikely that results will be made public by DfES until the summer of 2003. It is hoped that it will be possible to present further details of the survey at the ALM-10 conference.

In addition to the proportions at each of the five levels, some measure of relative difficulty of different topic areas - and hence of spiky profiles - may be obtainable. Certainly such spiky profiles are evident in individuals' performances to date. In general terms, there appear to be three categories of respondent, - the very small proportion of respondents who remained at or below Entry Level 1, the much larger group who were at Level 1 or 2 throughout, and the majority who found some topics hard and others much easier. Measures of relative difficulty of different topics may possibly be made, but these are compounded by questions to do with the facility of individual items and too much significance should probably not be given to them.

Further research

Although the multiple-choice style of presentation enabled the survey to be adaptive and greatly facilitated the collection and future analysis of the resulting data, the style will only produce limited information concerning individuals' numeracy capabilities, personal techniques and understandings.

Several possible further research projects immediately suggest themselves. These include presenting the items as short response items to a much smaller sample of adults, then recording and analysing the responses in order to gain insight into methods used to tackle individual items, and then to make comparisons with the multiple choice versions of the items. The same multiple-choice stems could also be re-used, but with alternative distractors.

Acknowledgements

In presenting this paper, the author would like to acknowledge with gratitude the contributions to the project from Linda North and Lynne Tranter in item writing, Iain Cummings and John Winkley (both of BTL), in software design and Dr Jenny Tuson (CDELL) and Joel Williams (BMRB) in data analysis.

References

Basic Skills Agency (2001) Adult Numeracy Core Curriculum , London, Basic Skills Agency.

Coben D., O'Donoghue J., and FitzSimons G. (2000) Perspectives on Adults Learning Mathematics: Research and Practice , Dordrecht, Kluwer

DfEE (1999) A fresh start: Improving numeracy and literacy , London, Department for Education and Employment (now DfES)

Ekinsmyth C. and Bynner J. (1994) The Basic Skills of Young Adults: some findings from the 1970 British Cohort Study , London, Adult Literacy and Basic Skills Unit.

OECD/UNESCO (1994-5 et seq.) International Adult Literacy Survey – IALS. Organisation of Economic Development (OECD), Eurostat, and UNESCO Paris

Qualifications and Curriculum Authority (2000) Key Skills units Levels 1 – 3 . London, QCA

van den Heuven-Panhuizen M. (1994) New chances for paper and pencil tests. Proceedings of 45th CIEAEM meeting , Cagliari, Italy

van den Heuven-Panhuizen M. (1996) Assessment and realistic mathematics education , Freudenthal Institute, Utrecht.

PISA (2001) Draft Framework for the PISA 2003 mathematics assessment , August 2001 Nijmegen, NL, Programme for International Student Assessment.

Disclaimer

This paper presents the views of the author only.

Figure A1 – algorithm for progression between steps

