Assessment of 5- to 14-year-old children's mathematical progress in schools in Wales

Llywodraeth Cymru
Welsh Government

The purpose of this rapid research study was to review what works in terms of numeracy testing for children aged 5- to 14-years-old. In order to do this, both current practice and the research literature were considered by the authors. Based on this exploration, the authors provide a set of recommendations for the planned national numeracy test in Wales.

Findings and recommendations:

• The authors found that there are four tests currently in use across Wales to test children's mathematical ability.

• These tests are predominantly commercially available and widely-used.

• These tests vary, but in the main they focus on the testing of children's mathematical procedures.

• A more thorough assessment of children's mathematical ability would be provided by testing both mathematical procedures and mathematical reasoning.

• A new mathematical test should produce both standardised and age-equivalent scores.

• A new test should be rigorously tested to ensure that it is both a reliable and valid measure.

Dadansoddi ar gyfer Polisi
YMCHWIL GYMDEITHASOL Y LLYWODRAETH
GSR
GOVERNMENT SOCIAL RESEARCH
Analysis for Policy

Introduction

This research summary is about the existing arrangements for assessing children's numeracy in schools in Wales, and also about how to replace these with a new All Wales Numeracy Test. By numeracy we mean: competence and confidence in using numbers to represent quantities and also the skills in reasoning which people can use in order to understand relations between quantities, to model problem situations, to make quantitative predictions, and also to assess the probabilities of different possible events.

The first step in the research summary is to review and analyse the tests that are currently in use as measures of mathematical achievement in Welsh schools. The next part is a description of the possible form and contents of a good and effective mathematical achievement test for children between the ages of 5- and 14-years throughout Wales. The research summary ends with a set of recommendations about how to construct such a test.

Review and analysis of tests currently in use

- Different local authorities in Wales use one or more of four mathematical achievement tests to assess the progress made by school children in their area in mathematics.

- One of the four tests was constructed by a local authority. The other three are commercially available and widely-used achievement tests.

- Despite some striking differences between these four tests, they have in common the fact that they all test children's use of the *mathematical procedures*[1] that they learn at school.

- In procedural problems, it is always clear what mathematical move the child must make and the test item tests how well he or she can make it. The mathematical procedures tested in the four tests include counting,

carrying out additions, subtractions, multiplications and divisions, using fractions and percentages, and, in some of the tests, reading tables and graphs, and making judgments about symmetry.

- Only two of the four tests contain any serious attempt to test *children's mathematical reasoning*[2]. This is a matter of concern, since previous research suggests that children's reasoning abilities play at least as important a role in their mathematical learning as does their knowledge of mathematical procedures.

- Mathematical reasoning usually takes the form of working-out the underlying numerical or quantitative relations in a problem in order to solve that problem. For example, we count the following problem as a test of mathematical reasoning because on the surface the problem appears to be about

addition, whereas in fact it is most easily solved by a subtraction: *John had 7 sweets and then Mary gave him some more. Now he had 13 sweets. How many did Mary give him?*

- The two tests given to Welsh schoolchildren which do contain reasoning items provide a rather inadequate picture of the children's mathematical reasoning, since very few of the test's items test children's ability to reason about central mathematical principles such as cardinality[3] and the inverse relation between addition and subtraction.

- In fact, we know of no existing educational or psychometric test that does a good, systematic job of testing mathematical reasoning, though it should be reasonably easy to construct such a test, since researchers have devised many ingenious measures of

a broad range of mathematical reasoning.

- Another question raised by the existing tests in use in Wales is whether a new test should be a paper-and-pencil assessment or a computer test, in which the items are presented on the computer screen and the participants key in their responses on the computer's keyboard. Some of the existing tests in use in Wales do one thing, some the other.

- The evidence that we have suggests that it is better to give young children up to the age of 8-years just paper-and-pencil tests. Computer tests are a possible option for children of 9-years and more, but there are some relatively complex items for which even the oldest children need paper and pencil to help them through.

Description of a good and effective mathematical achievement test

1. Our first requirement for a good mathematical achievement test is that it should contain a set of items that tests children's ability to reason mathematically as well as a set of items that measures their procedural skills.

2. The nature of these items should reflect the changes in children's mathematical reasoning abilities and in their procedural skills that take place as they progress through school. For example, the tests given to younger children should deal mainly with counting in whole numbers and with simple additions and subtraction on the procedural side, and with part-whole and correspondence relations (e.g. one-to-one and one-to-many correspondence[4]) as well as the additive

4

composition of number[5] and the inverse relation between addition and subtraction in the reasoning items. Older children, in contrast, should be given more complex procedural problems, some of which should involve fractions and decimals, and the reasoning items should be about multiplicative as well more complex additive relations.

3. The test should also include geometrical items and items about chance and probability. Again some of these should be procedural items and others reasoning items.

4. These suggestions fit well with the plans that are currently being made for the numeracy matrix with the proposed National Literacy and National Numeracy Framework. The numeracy matrix will include teaching in the early school years about measurement, chance, and use of the number system: all these

topics involve procedures and types of reasoning that are an important part of children's mathematical learning.

Recommendations for an all Wales mathematics test

1. The new test should contain procedural items measuring how children carry out the mathematical procedures that they are taught about at school and items that test how well they reason about quantitative relations when they have to, in order to find the best way to solve the problem at hand.

2. The assessment should provide information about the children's learning about counting, whole numbers, rational numbers[6], additive and multiplicative relations, geometry and probability, and also about algebra if this is part of what they are being taught.

3. The test should produce separate scores for

procedural skills and for mathematical reasoning, as well as a total overall score. The test should also provide separate scores for the different aspects of mathematics represented in the test (e.g. whole and rational number problems, additive and multiplicative reasoning etc.). The tests should produce clear standardised and age-equivalent scores.

4. The children's performance should be measured mostly by the number of items that they answer correctly. However, the scores for some items would reflect the steps the children take to solve the problems, as well as the correctness of their final solutions.

5. For the younger children in Years 1, 2 and 3, the tests should be paper and pencil tests and the questions should be read out to the children by the adult in charge. This will remove the possibility that some of these children's mistakes might be due to an inability to read the questions adequately.

6. The older children should be given paper and pencil items as well as items administered through computers. The assessment of some aspects of procedural knowledge can be easily carried out in computer-administered tasks but children may need to use paper and pencil for items that demand more complex reasoning or involve several steps in problems solving.

7. The test should be piloted, and its reliability and validity[7] tested, on a fairly large and representative sample of children. The test's validity should be assessed by relating the children's scores to an independent measure of their mathematical progress.

8. It should also be possible to look at the relation between

the children's scores in the mathematics assessment and their achievement in other subjects. One important link to make is with science; since much of children's scientific learning involves proportional concepts. The mathematics that children learn is also relevant to aspects of geography (e.g. the use of co-ordinates, the concepts of population density and income *per capita*) and history (e.g. time relations and also the question of the randomness and the probability of historical events). Since the framework for a new curriculum contains plans for emphasising these connections between numeracy and other subjects, it would be valuable to know how well children's mathematical learning prepares them for these related subjects as well.

Glossary of terms

1. Mathematical procedures: the steps taken to carry out a particular kind of calculation. Some of these procedures (like long division) are based on algorithms taught in the classroom.

2. Mathematical reasoning: logical reasoning about relations between quantities based on the knowledge of certain mathematical principles.

3. Cardinality: a basic property of number; refers to the fact that if the items in one set can be paired one-by-one with the items in another set, and neither has any items left over, then the sets are equivalent and so have the same cardinality.

4. One-to-one and one-to-many correspondence: the items in any two sets with the same cardinal number can be placed in one-to-one correspondence. For every item in one set there is an equivalent item in the other set (like two ranks of soldiers). This is a relative cue for cardinal number: one can judge that two sets are equal because they are in one-to-one correspondence, without knowing their actual number.

5. Composition of number: the principle that all numbers are made up of combinations of other numbers: e.g. 9 is a

combination of 6 and 3, and of 5 and 4, and of 6 and 2 and 1 (this also applies to negative numbers: e.g. -9 is a combination of -6 and -3).

6. Rational numbers: represent quantities by a relation between two numbers. The sets of rational numbers includes ordinary fractions like $3/4$ and decimals like 0.75 (but a whole number can also be represented as a rational number, as in $2/2$ or $4/2$).

7. Reliability and Validity: reliability: educational and psychometric tests are required to be reliable in two ways: 1) test-retest reliability: if the test is taken twice by the same people, there should be a strong correlation between their scores on the first and second occasion 2) inter-item reliability: the different items in the test should correlate significantly with the total score. Validity: a test is valid when it measures what it claims it measures. Usually the validation of an educational or a psychometric test is carried out by relating the participants' scores in the test to some external criterion. The usual validity criterion for mathematics tests given to children is their results in a school or state exam in mathematics or some other measure of their progress in mathematics at school. If the scores on the test predict these external measures it can be judged as a valid test.

Authors: Peter Bryant and Terezinha Nunes, Department of Education, University of Oxford