Assessing Number Sense: Collaborative Initiatives in Australia, United States, Sweden and Taiwan

Alistair McIntosh, Jack Bana & Brian Farrell
Edith Cowan University

Group tests of Number Sense were devised and administered to students aged 8, 10, 12 and 14 in Perth, Western Australia and in Missouri. The development and structure of the tests are described, and some results are presented. Research building on this work and conducted in Sweden and Taiwan is described, and implications for research and teaching are discussed

Purpose of the Research

In 1995 we reported on research into the mental computation of school students in Australia, Japan and the United States (McIntosh, Bana & Farrell, 1995). We were keen to explore these students' understandings of number more widely, and also to investigate the extent to which students' number sense could be meaningfully assessed by group administered paper and pencil tests.

We therefore made an analysis of number sense based on our own understandings and the descriptions and analyses of others and used this to structure tests which were administered to the same cohorts of Australian and American students as had been involved in the mental computation research.

Swedish colleagues at the University of Goteborg became interested in the work and collaborated with us in administering revised versions of the tests to cohorts of students of similar ages in Goteborg. A Taiwanese postgraduate student at the University of Columbia adapted some items and administered them to students in Taiwan, resulting in some valuable insights into the mathematical understandings of Taiwanese students.

The aim was not to produce definitive answers but to probe ways of exploring the very difficult area of student understandings and to acquire a better understanding of what number sense means and how it is exhibited in detectable and describable ways.

What Is Number Sense?

Number sense is not the same as numeracy, though it obviously overlaps with it. In today's discourse numeracy is generally understood to refer to the mathematics that one needs in order to deal with everyday life, and it includes more than number. For example, the definition adopted by the Literacy and Numeracy Benchmarking Taskforce in Australia begins: 'Numeracy is the effective use of mathematics to meet the general demands of life at school and at home, in paid work, and for participation in community and civic life'.

We take number sense to refer to a person's general understanding of number and operations along with the ability and inclination to use this understanding in flexible ways to make mathematical judgments and to develop useful and efficient strategies for managing numerical situations (for more detailed discussions of the components of number sense see McIntosh, Reys & Reys, 1992; Resnick, 1989; Sowder & Schappelle, 1989; Sowder, 1992; Willis, 1990).

Methodology

Sample

Students from two school districts (mid-sized to large school districts near a university) participated in the study, one in Perth, Western Australia, and one in Columbia, Missouri. Schools, then classes within schools, from the districts were randomly selected for participation. The number of students participating at each age level within each country ranged from 110 to 160 students. In both countries students aged 8, 10, 12 and 14 participated in the study. Some of these items together with some new items were presented to students aged 10 and 14 in Goteborg, Sweden. A few items were also adapted and used with students aged 12 and 14 in Taiwan.

Instruments

McIntosh, Reys, & Reys (1992) developed a number sense framework based on studying and reflecting on the literature associated with the topic. From the framework six strands were identified (two each from the three major categories of the framework). The six strands represent major components of our definition and understanding of number sense. The strands are:

- Understanding of the meaning and size of numbers.
- Understanding and use of equivalent representations of numbers.
- Understanding the meaning and effect of operations.
- Understanding and use of equivalent expressions.
- Computing and counting strategies.
- Measurement benchmarks.

The framework and identification of these six strands provided a structure for the development of number sense assessment items. After identifying the six number sense strands, a search and review of available assessment instruments was conducted to identify items which might provide a useful 'window' into the thinking of students as related to the various strands. The ideas of Markovitz & Sowder (1994) and Cramer, Post, & Currier (1993) were particularly useful in identifying items which mirrored elements of the framework. In addition, new items were developed to reflect various components of number sense. Items were constructed which were appropriate to the age level in terms of conceptual knowledge and which framed questions in non-routine environments in order to elicit strategy generation based on understanding rather than strategy recall based on familiarity with problem-type.

Early versions of the items were used in interviews with students in the United States. These interviews provided a means to clarify wording and item format and to probe the thinking of individual students as they responded to items. Items which evoked thoughtful introspection (rather than immediate recall of a rule or known fact) and which relied on conceptual understanding, were retained. On the other hand, items which could be answered correctly without indication of understanding were dropped.

Some items were organized into a format which allowed for a group-administered number sense test. Slightly different forms of the test were used in the two countries, but most items were identical. The number of items contained on each number sense test varied by grade level and country (ranging from 30 to 45 items).

The Number Sense Tests were administered during one class period at the beginning of the fourth quarter of the school year (March in the US and Sweden, August in Australia) and at mid-year in Taiwan. Items were read to the 8-year-old students. For all ages, a pacing scheme was used in the administration of the test. Specifically, the instructions were read and students were asked to spend no more than 30 seconds on each item. To encourage students to adhere to the suggested pace, the item number was announced each 30 seconds (for example, the administrator said, "You should be looking at item number three about now."). The pacing scheme was utilised to encourage children to answer each item based on reflection rather than computation and to prompt them to move along in the test rather than spend a lot of time on any one item. We found that this pacing scheme was useful in moving students through the various items and also conveyed the notion that computation proficiency was not the goal. Students wrote responses directly on the test pages. Total administration time was 30 - 40 minutes.

Some Results

To give the flavour of the results each of the six components will be described in more detail and the results of one item associated with each will be discussed. A full account of all items and results is given in McIntosh et al (1997).

1. Understanding of the meaning and size of numbers (Number Concepts).

Understanding of the base 10 number system (whole numbers, fractions, and decimals) includes patterns and place value which provide clues to the meaning/size of a number (e.g., $\frac{5}{6}$ is a fraction less than one, it is close to one because of the relationship between the numerator and denominator OR 1000 is a large number if you are referring to the population of a school but a small number if you are referring to the population of a town). It could involve relating and/or comparing numbers to standard or personal benchmarks. It includes comparing the relative size of numbers within a single representational form.

Table 1
Sample item (Ages 8 and 10)

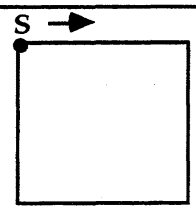
If I have \$378 in my savings account and withdraw all my money, how many 10-dollar notes would the bank be willing to give me?	________________

51 percent of the Australian 8-year-olds and 70 percent of the 10-year-olds answered this correctly, as did 53 percent of the Swedish 10-year-olds. The item differentiated well between quintiles. Percentages of Australian 8-year-olds answering correctly by quintile were respectively 88, 59, 36, 53, 18.

2. Understanding and use of equivalent forms and representations of numbers (Multiple Representations).

This strand includes recognition that numbers take many different numerical and representational forms (e.g. fraction as a decimal, a whole number in expanded form, or a decimal on a number line) and can be thought about and manipulated in many ways to benefit a particular purpose. It also includes the ability to identify and/or reformulate numbers to produce an equivalent form: the use of decomposition and recomposition to reformulate numbers for ease in processing; relating and/or comparing size of number(s) to a physical referent (e.g. collection of items, shaded region, or position on a number line); and crossing among various representational forms.

Table 2
Sample item (Ages 10 and 12)

You are going to walk once around a square-shaped field. You start at the corner marked S and move in the direction shown by the arrow. Mark with an X where you will be after $\frac{1}{3}$ of your walk.	

One quarter of the American and one half of the Australian 12-year-olds placed a cross within acceptable limits.

3. Understanding the meaning and effect of operations (Effect of Operations).

This strand includes understanding the meaning and effect of an operation either generally or as it relates to a certain set of numbers (e.g., Division means breaking a number into a specified number of equivalent subgroups OR Multiplying by a number less than 1 produces a product less than the other factor). It includes judging the reasonableness of a result based on understanding the numbers and operations being employed.

Table 3
Sample item (Ages 12 and 14)

Without calculating the exact answer, circle the best estimate for: 87 x 0.09.	A. a lot less than 87 B. a little less than 87 C. a little more than 87 D. a lot more than 87

The percentages of American, Swedish and Australian 14-year-olds answering this correctly were 50, 65 and 82 respectively.

4. Understanding and use of equivalent expressions (Equivalent Expressions).

This strand includes the translation of expressions to equivalent forms, often to re-evaluate and/or more efficiently process a computation. It includes understanding and use of arithmetic properties (commutativity, associativity, distributivity) to simplify expressions and to develop solution strategies (e.g. the use of the distributive property to multiply 7 x 52).

Table 4
Sample item (Age 14)

Write '*is greater than*', '*is equal to*' or '*is less than*' to make this a true statement:	$456 \div 8$ ________________ $456 \times \frac{1}{8}$

The percentages of American, Swedish and Australian 14-year-olds answering this correctly were 44, 46 and 48 respectively.

5. Computing and Counting Strategies

This strand involves applying various number sense components previously described in the formulation and implemention of a solution process to a counting or computational (estimation, mental computation, paper/pencil, calculator) situation (e.g. Is 29 x 38 more or less than 400? OR: How many birds do you estimate there are in this picture?)

Table 5
Sample item (Ages 8, 10, 12 and 14)

About how many days have you lived? (Circle the nearest answer.)	A	300
	B	3000
	C	30 000
	D	300 000

The significant digit in this item varied between countries: Australia used 3, the United States 4 and Sweden 5; the most surprising feaure of these results, which are discussed in greater detail in McIntosh, Reys & Reys (1995), is that across all three countries and four age groups the success rate only varied between 19 and 55. Indeed the

Australian percentages correct for the four age groups were respectively 34, 35, 38 and 32. The reasons for this apparent stagnation merit closer scrutiny.

Measurement Benchmarks

This strand was only added at the time of the Swedish testing. It includes applying various number sense components previously described in the formulation and implemention of a solution process to a measuring situation. It requires an understanding and use of standard, non-standard and/or personal benchmark units of measure (e.g. A textbook weighs about a kilogram OR: The angle is slightly less than a right angle so it must be about 85 degrees). It involves measuring attributes such as mass, length, capacity, volume, time and angles.

Table 6
Sample item (Age 10, Swedish version)

How long is a bed for an adult? Circle the best answer	A. 50 cm B. 100 cm C. 200 cm D. 400 cm

Sixty-eight percent of the Swedish children answered this correctly, with a further 20 percent choosing 400 cm.

Developments in Sweden

Researchers at the University of Goteborg carried out similar testing with students aged 10 and 14 and some of the results are given above. However the major innovation was that they published the tests in the journal *Namnaren* together with some discussion of the meaning and importance of number sense and invited teachers to try the tests in their own classroom and to share results and their opinions on the testing and their perception of the relevance of number sense. Thus the exercise initiated a mini-national debate, together with professional development and the involvement of teachers in action research. The project is described in Emanuelsson, Johansson, Reys & Reys (1996).

Developments in Taiwan

Der Ching Yang, a Doctoral student at the University of Missouri-Columbia devised a twenty-item Written Computation Test (WCT) for students aged 12 and 14 and a forty-item Number Sense Test (NST). The first 20 items of the NST were developed to parallel the WCT Items. That is, the same numbers were involved but the format of the items related to these numbers was different. The two tests were administered on different occasions.

The research provided strong evidence that Taiwanese students perform at very different levels on written computation and number sense tests. Basically, the number sense performance level for both 12-year-old and 14-year-old students in Taiwan was poorer than their written computation performance. One example will serve to indicate the difference between their written computation performance and number sense ability. The percentages of the same cohorts of students giving each reponse on two parallel items, one from the WCT and the NST respectively, were as follows:

Table 7
Sample Items (Taiwan, age 12 and 14)

NST Item			WCT Item		
Without calculating an exact answer, circle the best estimate for $\frac{12}{13} + \frac{7}{8}$.			Calculate $\frac{12}{13} + \frac{7}{8}$		
	Gr.6	Gr.8		Gr.6	Gr.8
A. 1	10	20	Correct	61	63
B. 2 *	25	38	Incorrect	39	37
C. 19	36	14			
D. 21	16	12			
E. I don't know	10	16			
F. No response	3	0			

Whereas 61 percent of the 12-year-old students and 63 percent of the 14-year-old students could correctly calculate $\frac{12}{13} + \frac{7}{8}$, only 25 percent and 38 percent respectively of these same students gave 2 as the best estimate for the computation.

Discussion

The research supports the view that there is some identifiable field called number sense and that the number sense of students can, to a certain degree, be investigated and assessed through the use of group-administered written tests. However, often the research produced results whose main value was that they indicated a potentially rich field of inquiry which merited further exploration through structured interviews with individual students.

It was apparent that students had many deficiencies in understandings, particularly when fractions and decimals were involved.

There is scope for much further research, into more sensitive assessment items and procedures, into the performance of students and their thinking, and into ways of translating the implications of the research into effective classroom action.

All of this assumes even greater importance in view of the current moves at state and federal level in regard to the setting of numeracy benchmarks and associated assessment procedures which reveal yet again the tension between the expertise and aspirations of professional educators and those whose standpoint is more remote from the classroom and more simplistic.

References

Cramer, K., Post, T., & Currier, S. (1993). Learning and teaching ratio and proportion: Research implications. In D. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp. 159-178). New York: Macmillan.

Emanuelsson, G., Johansson, B., Reys, B.J. & Reys, R.E. (1996). *Assessing the Development of Number Sense: Using a Journal to Engage Swedish Teachers in Developmental Work.* Paper presented at the 1996 Annual Meeting of the American Educational Research Association, New York, April 1996.

Markovits, Z. & Sowder, J.T. (1994). Developing number sense: An intervention study in grade 7. *Journal for Research in Mathematics Education, 25*, 4-29.

McIntosh, A., Reys, B., Reys, R., Bana, J. & Farrell, B. (1997). *Number sense in school mathematics: Student performance in four countries.* Perth: MASTEC, Edith Cowan University.

McIntosh, A., Reys, B.J., & Reys, R.E. (1992). A Proposed framework for examining basic number sense. *For the Learning of Mathematics, 12*(3), 2-8.

McIntosh, A., Bana, J. & Farrell, B. (1995). Mental Computation in Australia, Japan and the United States. In *Galtha: Conference Proceedings of the Eighteenth Annual Conference of the Mathematics Education Research Group of Australasia (MERGA).* Darwin: MERGA

McIntosh, A.J., Reys, B.J., & Reys, R.E. (1995). How Old Are You? *Mathematics Teaching 153* , 5-7.

Resnick, L. B. (1989). Defining, assessing and teaching number sense. In J. Sowder & B. Schappelle (Eds.), *Establishing foundations for research on number sense and related topics: Report of a conference.* San Diego, CA: San Diego State University, Center for Research in Mathematics and Science Education.

Sowder, J. T. & Schappelle, B. P. (Eds.). (1989 *Establishing foundations for research on number sense and related topics: Report of a conference*. San Diego, CA: San Diego State University, Center for Research in Mathematics and Science Education.

Sowder, J. T. (1992). Estimation and related topics. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp.371-389). New York: Macmillan.

Willis, S. (Ed). (1990). *Being numerate: What counts?* Hawthorn, Victoria: Australian Council for Educational Research.