What does Effective Teaching for Numeracy Look Like? The Design of an Observation Schedule

Tracey Muir
University of Tasmania
tracey.muir@utas.edu.au

This paper reports on the development and refinement of an observation schedule designed to evaluate effective teaching for numeracy and to serve as a starting point for teacher self-reflection. This paper summarises the key findings from the literature as to what constitutes effective numeracy teaching and documents the results of a pilot study designed to test the usefulness of the instrument. The results indicate that the instrument is quite comprehensive in terms of its coverage of effective numeracy teaching indicators and may prove useful to other researchers involved in documenting classroom practice.

The term 'numeracy' was first coined in 1959 in the Crowther Report (Department of Education and the Arts, 1995) and later in 1982, where the Cockcroft Report (Cockcroft, 1982) referred to it as implying an 'at-homeness' with numbers, and an ability to make use of mathematical skills which enable an individual to cope with the practical demands of everyday life. In Tasmania, the current definition of numeracy used by and referred to by Tasmanian teachers is the one which describes the Being numerate element in the state's Essential Learnings Curriculum as: "Understands and has the confidence and disposition to meet the demands of life" (Department of Education, Tasmania, 2002, p. 5). This definition is consistent with the popular definitions identified in the literature (e.g., Steen, 1997; Johnston, 1994; Cockcroft, 1982) and is used as the referent in this paper as the study was conducted in a Tasmanian context.

The success of particular Asian countries in the TIMSS studies conducted in 1995, 1998 and 2003 sparked an interest in describing commonalities between classrooms of particularly effective countries (Clarke & Clarke, 2002). Individual countries have also sought to identify what constituted effective primary teaching of mathematics. In 1997, Askew, Brown, Rhodes, Johnson and Wiliam identified effective teachers of numeracy in a range of schools in the UK by looking at mean test scores of students over time. Data were collected on over 2000 pupils and evidence was gathered from a sample of ninety teachers. While their definition of numeracy differed slightly from that used in Australia in that it included number understanding and skills out of context as well as applications, they were able to identify characteristics which distinguished highly effective teachers of numeracy from other teachers. While according to Stephens (2000) there appear to be no comparable studies of numeracy undertaken on a similar scale in other countries, Askew et al.'s (1997) findings have been supported by other research in this area (e.g., Saunders, 2004; Smith & Geller, 2004; Marshall, 2003; NSW Department of Education and Training, 2003; Clarke & Clarke, 2002). A synthesis of the literature revealed some commonalities with respect to effective teaching for numeracy and these commonalities were used to provide the theoretical framework for this study. Key findings indicate that effective teaching for numeracy includes:

- A focus on important mathematical ideas
- The use of teaching approaches which foster connections between both different areas of mathematics and previous mathematical experiences

- Encouragement of purposeful discussion through the use of question types to probe and challenge children's thinking and reasoning and encouraging children to explain their mathematical thinking
- Knowledge and awareness of conceptual connections between the areas which they taught of the primary mathematics curriculum and confidence in their own knowledge of mathematics
- A particular emphasis on the development of mental computation.

Although not all studies identified the importance of mental computation explicitly, numeracy professional learning in Tasmania has recently focused on this aspect and thus its inclusion is particularly relevant to the Tasmanian context. A belief that more interactive forms of whole class teaching can play a vital role in raising literacy and numeracy standards (Hardman, Smith, Mroz & Wall, 2003) has also resulted in the advocating of a lesson approach that involves an oral and mental starter, a main teaching activity and a reflective plenary. This paper looks at the conversations involved in the first part of the lesson. Interestingly, the Askew et al. (1997) study found no common form of classroom arrangement between the case-study teachers in the highly effective group, despite this return to the concept of whole class teaching and the three-part lesson structure that was recommended by the UK National Numeracy Strategy (DfEE, 1999).

Effective Teaching Behaviours

In order to be an effective teacher of numeracy, the literature identified that teachers must be able to focus on key mathematical ideas, make connections between different mathematical ideas, draw out students' understandings and use their explanations as teaching points and challenge all students to reach their mathematical potential. These characteristics, however, involve general principles that may not be readily identifiable and do not portray specific teacher behaviours or describe what happens in effective classroom numeracy practice. While some of the aforementioned studies have identified specific strategies or practices used by these teachers, the author wished to design an observation schedule that was detailed enough to make explicit the types of teaching behaviours that constitute effective teaching for numeracy and furthermore, to use the students' responses and/or written work as evidence of how effective these behaviours were in helping students understand.

Observation of classroom discourse over many years has revealed it typically to be heavily dominated by teacher talk and even when questioning students, teachers tended to dominate the discourse (Galton, Hargreaves, Comber, Wall & Pell, 1999). Traditional classroom discourse has involved an exchange of 'initiation – response – feedback', providing little opportunity for students to engage in extended responses to express and evaluate ideas of their own (Tanner, Jones, Kennewell & Beauchamp, 2005). Tanner et al. (2005) cite research by Smith et al. that suggests that the use of questioning to scaffold students' learning is underexploited, with only 25% of questions designed to encourage students to think more deeply about their ideas.

Questions have also been used traditionally as a means of keeping students on task and as a way of directing the attention of the class. The responses are typically extracted from a number of individual children and rarely does the teacher have an extended exchange with a single student, involving a second, third, fourth or even fifth follow-up question (Wiliam, 1999). Hardman et al. (2003) also found that teacher questioning rarely went beyond the recall and clarification of information, with the teacher exercising close control over the nature and pace of the lesson. They particularly noted the use of the teachers' response to students' answers by observing the occurrence of probing questions (where the teacher stayed with the same child to ask further questions) and uptake questions (where the teacher incorporated a pupil's answer into a subsequent question) (Hardman, et al., 2003). These particular aspects were taken into account by this researcher when designing the observation schedule.

The importance of using questioning as a scaffolding tool was emphasised in a report detailing research into numeracy teaching approaches in primary schools (Commonwealth of Australia, 2004). The report identified twelve interaction or scaffolding practices which were "engaged in by teachers to support student learning that might ultimately be removed when students are able to 'stand alone' in respect to what they have learned" (p. iv). Along with questioning, another practice identified was modelling and the report defined this practice as "demonstrating, directing, instructing, showing, telling, funnelling, naming, labelling, explaining" (p. v). The researcher particularly selected modelling from these practices in order to document how teachers chose to represent concepts, their choice of examples and use of concrete materials or representations — specific behaviours which could serve to illustrate characteristics of effective numeracy teaching referred to earlier.

The Observation Schedule

An extensive search of the literature failed to identify an existing observation schedule that would meet the needs of this research. Observation schedules that were used in the studies previously identified were limited in that they tended to focus on general teacher behaviours, rather than explicit behaviours and similar studies often reported results, with insubstantial information provided about what the instruments actually 'looked like'. Detailed observation studies, such as those conducted by Galton, Simon and Croll (1980), Galton, et al. (1999) and Stodolsky (1988), influenced the design of the instrument, along with aspects of the observation schedules used by The Commonwealth of Australia (2004), Clarke and Clarke (2002), Hill (2000) and Rhodes, Swain, Coben and Brown (2004).

The observation schedule design focused on documenting the frequency and quality of teacher behaviours in terms of making connections, choice of examples, questioning and responding to students' answers, previously identified as being important elements of an effective numeracy teacher. The schedule also provided for documentation of mental computation opportunities and tracking of students' responses. A rating scale was devised in order to map the quality of the exchanges that would occur (see Table 1).

The researcher was also interested in the exchanges that occurred between teachers and their students and whether or not there was evidence of probing and uptake questions as identified by Hardman et al. (2003). This was also incorporated into the schedule, with its codes identified in Table 2.

Student responses were also deemed to be an important aspect of the discourse and an indicator of the effectiveness of the teacher's instruction. A summary of the way student responses were coded is outlined in Table 3.

Table 1
Rating scale for teacher behaviours

Teacher behaviour	NA	0	1	2
Making Connections (MC)	No connections necessary	Opportunity for connections, but none provided	Connection made, but did not serve to enhance understanding	Relevant and purposeful connection made
Choice of examples (CE)	No example necessary	No example provided (but one would have been helpful)	Poor example provided	Good example provided
Use of concrete materials (CM)	Use of concrete materials not necessary	Concrete materials could have been useful but not utilized	Materials used to obtain answers	Materials used as an integral/sense-making tool
Mental Computation (Ment. Comp.)	Use of mental computation not applicable	Mental computation not encouraged	Mental computation called for, but not focus	Explicit focus on mental computation
Use of open-ended questions			Calls for explanation (may begin with how/why)	Calls for justification, generalization or seeking of alternatives (may begin with 'what if?'

Table 2
Teacher response codes

Tr. response	Description
Generic feedback (GF)	May include praise or generic response, such as 'good answer'
Specific feedback (SF)	Draws attention to mathematics inherent in answer (e.g., good to see you using the strategy of bridging 10 to add those numbers)
Probe	Teacher stays with the same student to ask further questions
Uptake	Teacher incorporates a student's answer into a subsequent question
Repeat answer (RA)	Teacher repeats student's answer verbatim
Repeat question (RQ)	Teacher repeats or rephrases question
Clarifies answer (CA)	Teacher rephrases student's response

Table 3
Student response codes and meanings

Student response	Description
No response (NR)	No answer is volunteered to the teacher's question
Chorus	Whole class responds orally to the teacher's question
Individual student (S1)	Denotes which student answers (enables link with probing questions by teacher)
Correct (C)	Correct response given
Partially correct (PC)	Partially correct response or incomplete response given
Incorrect (I)	Incorrect response given
Sharing (S)	Requires student to give information (e.g., What can you tell me about money?)
Explanation (E)	Requires student to explain answer or strategy (e.g., How did you work it out?)
Justification (J)	Requires student to explain *why* they chose a particular process or strategy
Question (Q)	Student answers the teacher's question with another question

Provision was made to record episode numbers, with an 'episode' signifying the onset of a new teacher behaviour (e.g., when the teacher asked a different question), and to describe the behaviour of the teacher and students during each episode, therefore allowing for both quantitative and qualitative measures to be incorporated. Following the initial design of the observation schedule, a pilot study, focussing on the practice of one teacher, was conducted to test the application of the instrument.

The Pilot Study

The research was conducted using naturalistic inquiry (Lincoln & Guba, 1985) from an ethnographic perspective, as the intention was to create as vivid a reconstruction as possible of teachers' classroom practice (LeCompte & Preissle, 1993). The study was also consistent with Hitchcock and Hughes' (1989, as cited in Cohen, Manion and Morrison, 2000) claim that ethnography involves the production of a list of features common to a particular group and the description and analysis of patterns of social interaction. Ethnographic inquiry has also been particularly suitable for identifying and understanding social and cultural norms in mathematics classrooms (Walshaw & Anthony, 2004). Data were collected for the pilot study through observation, although other ethnographic aspects were incorporated into the major study, including interviews and participation observation.

The teacher for the pilot study was chosen using opportunity sampling and was known to the researcher. Although the main study will focus on upper primary teachers, circumstances and time constraints led to the selection of this teacher and her grade 2/3 class. The teacher, Susie, had taught a variety of grades and at a variety of schools throughout her fifteen-year career. She occupied a senior position in her school, which catered for students in grades k-10. There were 23 students in the class.

The researcher visited the classroom of the teacher on one occasion and video taped the introductory and plenary sessions which involved whole group discussions. The videoing was necessary to fully capture the exchanges that occurred and to enable the researcher to more accurately analyse the teaching behaviours using the comprehensive observation schedule. Arrangements were made for a third person to videotape segments of lessons in the previous week in an attempt to minimise the impact of the video's presence on the students. The teacher was asked to document the aims of the lesson, but no formal lesson plan was produced.

The researcher fully transcribed the videotape within hours of the observation, as some of the sound quality on the video made it difficult to hear some students' responses. The transcript was cross-checked with the teacher to ensure optimum accuracy. The researcher then viewed the video, using the observation schedule to code behaviours and referring to the transcript when necessary. As this was the first time the researcher had used the instrument, frequent pausing of the video was required in order to document what was happening in the discourse. Some details were amended and some categories were added which resulted in the rating scales earlier documented. After the third viewing of the video footage, the researcher was able to use the observation schedule to record what was happening in 'real time'. The video footage also allowed the researcher to take note of wait times and record the amount of teacher talk and the length of students' responses.

Results and Discussion

The results and discussion pertain to the introductory part of the lesson and examples are provided to illustrate some of the coding that occurred. Field notes and video evidence indicated that the teacher's questions were predominately focused on the mathematical task and the discourse characteristically took the form of the teacher asking a question and allowing for approximately five or six student responses before asking another question. The students were attentive and all students had the opportunity to contribute.

Student response codes allowed all student exchanges to be coded. The inclusion of a category ‘general teacher talk’ would also have allowed for all teacher behaviours to be coded. There was only one example of ‘choice of example’ included in the discussion and that was presented up front and rated a ‘2’ in that it was a ‘good example provided’. Connections made by the teacher predominately focused on links with previous learning and these were also rated ‘2’ as they were deemed to be relevant and purposeful. The teacher did not make direct use of concrete materials, although these were referred to in the course of the discussion. There were opportunities for mental computation with a significant part of the discussion focusing on this aspect. A total of 13 open-ended questions were asked; six of these were classified as ‘Open 1’, while the remainder were non-mathematical. No ‘Open 2’ questions were asked. These findings are consistent with the research which suggests that the use of questioning to scaffold students’ learning is underexploited (Tanner et al., 2005). A frequency count of teacher and student behaviours revealed some interesting trends and these are presented in Tables 4 and 5.

Table 4 shows that there were a high number of instances of generic feedback, compared with only two examples of specific feedback. Examples of generic feedback included comments such as “Well done” and “Good answer”. An example of specific feedback was “Good to see you counting on in 5s to obtain the answer”. There were 18 examples of probing questions with the majority of these occurring in response to students’ partially correct responses. For example, when one student responded to the question, “How would you go about adding $1.50 and $1.25?” with “I would add the two one dollar coins first”, the teacher continued to ask a series of probing questions of the same student in order to eventually produce the correct response. Rarely did the teacher stay with the same student when a correct answer was produced, but would instead move on to another class member for their response.

Table 4
Teacher response codes

Tr. response	Number of instances
Generic feedback (GF)	12
Specific feedback (SF)	2
Probe	18
Uptake	4
Repeat answer (RA)	18
Repeat question (RQ)	6
Clarifies answer (CA)	5

Table 5 shows the number of instances related to the categorized student responses. It was sometimes difficult to distinguish between sharing and explaining, but answers were classified as ‘sharing’ in response to questions such as, “Who can tell me what they know about money?” and “What could you use if you couldn’t work it out in your head?” Examples of explanations included:

> Well I’d add the 2 ones together, the one dollars and I’d go 50, and then I’d plus 10 on from 50 and then I’d add another 10 and then 5

> You'd take the two one dollar coins, which are two dollars and then you'd add up the cents — you add them the 50 cents and the 25 cents, but instead of adding on the 25, you add on 20 then add on the 5.

There were no instances of answers involving justification. The video footage also allowed the researcher to examine the lengths of students' responses, 80% of which were less than ten words. This is consistent with the research cited earlier by Tanner et al. (2005).

Table 5
Student response codes and meanings

Student response	Number of instances
No response (NR)	6
Chorus	4
Individual student (S1):	
Correct (C)	22
Partially correct (PC)	12
Incorrect (I)	4
Sharing (S)	25
Explanation (E)	17
Justification (J)	0
Question (Q)	4

More testing of the observation schedule needs to be undertaken to more accurately determine its reliability and validity. At this stage, only the researcher has used the schedule with repeated viewings producing consistent coding. In terms of validity, the observation schedule does allow for documentation of the behaviours described and all interactions could be coded. The results indicated that the observation schedule could prove to be a useful starting point for teachers to reflect on their practice through highlighting current behaviours, considering how these vary from the behaviours of highly effective teachers, and suggesting possible behaviours to incorporate into their numeracy discourse. In this particular case, completion of the schedule revealed the absence or under utilisation of some of the identified effective characteristics of numeracy teaching; Susie was surprised, for example, by the high number of 'lower level' open-ended questions and determined that some of these could be rephrased to require more justification of answers.

Conclusions

The observation schedule proved useful in determining the presence of effective characteristics of numeracy teaching as indicated in the literature and in raising the teacher's awareness of aspects of her own practice. Through consideration of these factors and viewing of video footage in a reflective manner teachers may determine ways to increase the effectiveness of their discourse. The comprehensive nature of the schedule necessitated the incorporation of video footage and the schedule would need to be modified to accommodate situations when video footage was not feasible. While the observation schedule has been trialled in only a limited way, it is hoped that its documentation may provide other researchers with a useful tool for classroom observations.

References

Askew, M., Brown, M., Rhodes, V., Johnson, D., & Wiliam, D. (1997). *Effective teachers of numeracy*. London: School of Education, King's College.

Clarke, D., & Clarke, B. (2002). Challenging and effective teaching in junior primary mathematics: What does it look like? In M. Goos & T. Spencer (Eds.), *Proceedings of the 19th Biennial Conference of AAMT* (pp. 309-318). Brisbane: AAMT.

Cockcroft, W. C. (1982). Mathematics Counts: Report of the Committee of Inquiry into the Teaching of Mathematics in Schools. London.

Cohen, L., Manion, L. & Morrison, K. (2000). *Research methods in education (5th Ed)*. London: Routledge Falmer.

Commonwealth of Australia. (2004). *Researching numeracy teaching approaches in primary schools*. Victoria.

Department for Education and Employment. (1999). *The National Curriculum: Mathematics*. London: DfEE.

Department of Education and the Arts, Tasmania. (1995). *Numerate students - numerate adults*. Hobart, Tasmania: Department of Education and the Arts.

Department of Education, Tasmania. (2002). *Essential Learnings framework 1*. Tasmania: Department of Education.

Galton, M. J., Hargreaves, L., Comber, C., Wall, D. & Pell, A. (1999). *Inside the primary classroom: 20 years on.* London, New York: Routledge.

Galton. M., Simon, B. & Croll, P. (1980). *Inside the primary classroom*. London: Routledge & Kegan Paul Ltd.

Hardman, F., Smith, F., Mroz, M., & Wall, K. (2003). *Interactive whole class teaching in the national literacy and numeracy strategies.* Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh.

Hill, P. (2000). *Numeracy education: what do we know and what can we learn from the literacy experience.* Canberra: ACT: AAMT.

Johnston, B. (1994). Critical Numeracy. *Fine Print, 16*(4), 32-36.

LeCompte, M., & Preissle, J. E. (1993). *Ethnography and qualitative design in educational research (2nd ed)*. London: Academic Press Ltd.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. London: Sage.

Marshall, J. (2003). Math wars: Taking sides. *Phi Delta Kappan, November 2003*, 193-200.

New South Wales Department of Education and Training. (2003). *Quality teaching in NSW public schools A classroom practice guide*: Professional support and curriculum directorate.

Rhodes, V., Swain, J., Coben, D. & Brown, M. (2004). *Researching teaching and learning of adult numeracy: designing a lesson observation schedule*. Retrieved November 12, 2005 from http://www.icme-organisers.dk/tsg06/Papers/SwainJ-23-Rhodes.pdf

Saunders, P. (2004). *Characteristics of effective teaching*. Retrieved April 18, 2005 from http://www.wmich.edu/teachlearn/new/char_effect_tch.html

Smith, K. S., & Geller, C. (2004). Essential principles of effective mathematics instruction: Methods to reach all students. *Preventing school failure, 48*(4), 22-29.

Steen, L. A. (1997). *Why numbers count: quantitative literacy for tomorrow's America*. New York: College Entrance Examination Board.

Stephens, M. (2000). *Identification and evaluation of teaching practices that enhance numeracy achievement*. Canberra.

Stodolsky, S. (1988). *The Subject Matters*. Chicago: University of Chicago.

Tanner, H., Jones, S., Kennewell, S., & Beauchamp, G. (2005). Interactive whole class teaching and interactive white boards. In P. Clarkson, A. Downton, D. Gronn, M. Horne, A. Mcdonough, R. Pierce & A. Roche (Eds.), *Building connections: Research, theory and practice, Proceedings of the 28th annual conference of the Mathematics Education Research Group of Australasia* (pp. 720-727). Melbourne: MERGA.

Walshaw, M., & Anthony, G. (2004). Research trends in mathematics education. In B. Perry, G. Anthony & C. Diezmann (Eds.), *Research in mathematics education in Australasia 2000-2003* (pp. 3-28). Sydney: MERGA.

Wiliam, D. (1999). Formative assessment in mathematics: Part 1: Rich questioning. *Equals, 5*(2), 15-18.