ASSESSING CHILDREN'S MATHEMATICAL THINKING IN PRACTICAL MODELLING ACTIVITIES

Howard Tanner and Sonia Jones, University of Wales Swansea, Wales, UK

This paper reports on research into the development of mathematical thinking skills using practical modelling activities. An action research group of twelve teachers from six secondary schools in Wales was created to investigate the use of mathematical modelling tasks in years seven and eight (11 and 12 years old). Twelve intervention classes were matched with twelve equivalent control classes as part of a quasi experiment to investigate the impact of such activities. The experimental design was based on the use of written pre-tests, post-tests and delayed tests. However, in addition to the written tests, 48 students - two from each class - engaged in mathematical modelling tasks during one to one semi-structured interviews with the researchers. The interviews used a form of *dynamic assessment* to analyse the development of students= mathematical thinking at each assessment point. This paper focuses on the interview data and reports on the extent to which some students proved to be naturally mindful and the extent to which metacognitive thinking might be taught in the early years of secondary school.

The thinking skills course Although based around a quasi-experimental design, the project had many of the characteristics of action research. The initial course materials were based on the activities devised during phase one of the Use and Practical Applications of Mathematics Project (PAMP) (Tanner & Jones, 1993b). However, it was intended that these activities would be further developed and new materials added. The teaching approaches were guided by those developed in phase one, but were also regarded as experimental and subject to development (Tanner & Jones, 1994a, p77-78).

Thus there were two strands to the course:- a structured series of cognitive challenges to stimulate the progressive evolution of key skills in the areas of strategy, logic and communication; - the use of teaching techniques which were intended to encourage the maturation of the metacognitive skills of planning, monitoring and evaluating.

The thinking skills targeted were metacognitive rather than cognitive (Brown, 1987; Gray, 1991). That is to say the course aimed to teach the processes rather than the content of mathematics through practical problem solving and modelling.

The research study planned to evaluate whether metacognitive skills had been taught successfully by testing for 'near transfer' (Shayer & Adey, 1992, p116) meaning that pupils who had followed the course would demonstrate improved performance in modelling situations which were similar in character to those used in teaching, but did not repeat the content of the lessons. The achievement of near transfer in this way would be non-trivial as research shows that even when pupils have strategic knowledge they may fail to apply it in problem isomorphs which have only slightly different outward appearances (eg: Gick and Holyoak, 1980; Tanner & Jones, 1994b). The contexts used in testing were physically different from those met in the intervention lessons.

The course activities were designed to encourage the development of a small number of general strategic or cognitive tools. Each activity or task was also targeted on at least one of the schemata of formal operations, eg: controlling variables, proportionality, correlation, probability.

A list of strategic skills was identified at the start of the materials. This list was not exhaustive or definitive, but was intended to aid in the grouping of activities to ensure that a variety of strategies was encountered in problem solving contexts to avoid an algorithmic approach or pseudo-problem solving. Strategies were not addressed separately in the activities - skill in comparing and selecting strategies was required. Activities were grouped to ensure that each group of activities was responsive to a small number of target strategies. It was intended that a pupil who had attempted an activity from each group would have encountered a wide range of strategies (Tanner & Jones, 1995a, p14-15). Metacognitive skills were not taught through the content of the materials but through the teaching approaches used. The teaching approaches employed were considered to be more significant than the activities chosen to provide contexts for learning. The teaching approaches utilized in the course were intended to encourage pupils to construct and evaluate their own strategies through discussion and debate (Tanner and Jones, 1995b; 2000).

The research design

This research has at its heart a quasi-experimental design involving pre-test, post-test and delayed-test of control and experimental groups which might be written as follows:

	Pre		Post	Delay	
	O1	X	O2	O3	12 experimental groups
	- -				
	O4		O5	O6	12 control groups
February			July	November	

It is not claimed that the twelve experimental groups all received an identical experience or 'experimental treatment'. In fact, the novelty of the approaches required and the materials used as context implied that the course experienced by pupils would vary considerably. In order to ensure sufficient internal validity, ongoing observation and analysis was undertaken of the way the course was being interpreted and experienced by teachers and pupils in the different social contexts which were developed in individual schools and classrooms.

The assessment instruments

Before the start of the course, tests and interviews were developed and trialled to assess pupils' cognitive and metacognitive knowledge and skills. Written tests were used in order to allow large scale testing of the control and intervention pupils according to the quasi-experimental design.

In addition to the written tests, techniques were also developed to assess pupils' metacognitive skills through the use of semi-structured interviews. Although the interviews may be considered to carry more face validity than the written assessments, they were very time consuming, and their use was consequently restricted to comparatively small numbers of pupils. However, their use gave the researchers closer insights into the developmental processes which were occurring.

Assessing metacognition

Metacognition is associated with awareness and control of one's own learning, (Brown, 1987). It includes an awareness of what one knows and does not know, the ability to predict the success of one's efforts (Royer et al, 1993), the planning, monitoring and evaluating of one's work (Gray, 1991), and an ability to reflect on the learning process and know what one has learned. It can be divided into passive knowledge and active skills. Observing essentially hidden metacognitive processes is far from easy, not least because people are adept at using small verbal or non-verbal cues to attempt to provide the responses which they think are expected. Several methods for eliciting information about thinking processes have been identified (Rowe, 1991) and a variety of direct and indirect approaches were used in the project to study students' metacognitive abilities.

Assessing the active metacognitive skills of planning, monitoring and evaluating

Metacognitive skills are inevitably contextualised when in use. Planning is necessarily purposive. The skills are also intricately related to each other, with monitoring and evaluating requiring a both a plan and a purpose. The active metacognitive skills of planning, monitoring and evaluating had been identified (Tanner & Jones, 1993a; 1994a,b) as underpinning practical mathematical modelling and so were measured in the context of solving practical problems requiring mathematical modelling. The assessments thus do not assess abstract generic skills directly, but estimate them from their application in a specific mathematical problem solving context. Novel problem solving contexts were chosen each time to ensure that more than the repetition of procedural knowledge was demanded.

The written assessment of active metacognitive skills

Metacognitive skills were assessed through a section in the written paper entitled 'Planning and doing an experiment'. In this section the students were required to apply their mathematical knowledge to solve a practical problem requiring mathematical modelling. They were tested on their use of mathematics in a novel situation. The contexts chosen were: Pre-test: Investigating a pendulum, Post-test: Toppling towers, Delayed test: Diving boards.

Two of these contexts now appear as tasks in the published course (Tanner and Jones, 1995a) but only the task from the pre-test was used in the course taught for the quasi-experiment. This was to ensure that the context used in assessment remained novel for both control and intervention pupils. The pre-test task is described below. Students were told that some string and a place to hang it from, a weight holder and some 20g weights, a tape measure and a stop watch were available. They were then asked to think of one interesting mathematical question to investigate using the equipment and to write down their plan under the four headings: My question; My plan; I would take these measurements; How I would present my results. Answers were assessed according to a set of criteria:

- the number of variables investigated eg: 'How long does it swing?' or 'I would compare swing with weight',
- whether variables were controlled,

- whether a relationship was sought and the quality of that relationship, eg: binary – 'long / short string versus time' or continuous – 'time measured for 20cm, 30cm, 40cm, etc.',
- the presentation of results eg: bar chart, ordered table, graph of ... against ..., seeking an equation or relationship.

The results of an imaginary experiment were presented and the students were invited to plot them on a graph, make a prediction, test the prediction against a formula and suggest how the results could have been made more accurate. The different problems used in the post- and delayed-tests allowed similar lines of development.

Interview based assessment of active metacognitive skills

Two pupils were interviewed from each intervention and control class. Each class teacher was asked to identify one pupil towards the top of their class and one towards the bottom. Interviews were conducted on a one to one basis between the university researchers and pupils whilst the pupil attempted to organize and conduct a mathematical investigation into a practical task. The same pupils were interviewed at each assessment point. Each interview lasted twenty to thirty minutes.

Students were assessed through a form of dynamic assessment, (cf: Feuerstein, 1979; Brown & Ferrara, 1985; Newman, Griffin & Cole, 1991). The researchers aimed to provide the minimum level of structure necessary for students to progress. The intention was to work in the student's 'zone of proximal development' (Vygotsky, 1978 p.86). Rather than observing students either succeed or fail in a task without intervention, the intention was to record how much help students required to make progress in a task.

Interviews followed a strict script which included settling down questions, instruction in how to use the equipment and a series of prompts to be used if students failed to progress. The interviewer had to make a judgement as to whether a prompt was needed to ensure progress. Interviews were tape recorded and transcribed. Assessments were made against specific criteria for levels of skill in planning, monitoring, evaluating and reflecting during the experiment. These assessments were then checked against transcripts. The written tests, interview scripts and assessment framework may be found in Tanner (1998).

Students were encouraged to think aloud during the task by such devices as:

> *Pretend that I'm your partner, but I'm not as clever as you. You have to explain things clearly so that I can understand what we are doing.*

For the pre-test interviews a simple pendulum was set up by the researcher in front of the student and then dismantled. Students were then asked to set up a similar arrangement for themselves. They were encouraged to keep talking throughout the experiment.

> *Talk to me as much as you can. I'm interested in all your ideas.*

Students were then encouraged to identify variables.

> *Your pendulum didn't have to be exactly the same as mine. What things can you think of which you might have changed?*

A series of prompts followed until sufficient variables were identified. They were then asked to hypothesize about which might affect time, using further prompts. After this they were asked to set up an experiment to investigate the pendulum. Marks were awarded for each level achieved in planning, monitoring, evaluating and reflecting. Marks were deducted for prompts given in each section. If prompts exceeded marks achieved, zero was awarded for that section.

An example of a criterion statement:

> 3 marks Shows evidence of planning to control variables and work systematically using binary logic, eg: times for long string and a short string.

An example of a prompt:

> Prompt 3 You said we could change ..... How could we test to see if it made a difference?

The script was trialled and developed through several different versions in one to one interviews in non-participating schools prior to the quasi-experiment.

The results

The reactions of the pupils to the interview tasks varied widely. Many children selected an approach to the task very quickly on encountering the situation and persisted with initial strategies in the face of mounting evidence of their failure. For example, on being presented with a pendulum to investigate, many pupils chose to time how long it would take a pendulum to stop. In spite of repeated prompts to move them into more successful strategies, many persisted and were still counting swings after 20 minutes. In the face of the obvious failure of this strategy others chose a point at which to 'cheat', grabbing the pendulum and claiming it had stopped. Others seemed to have a tendency to mindfulness and an inclination to stop and reflect. Several pupils learned from the assessment experience. One particularly mindful pupil from a control group began the second assessment by announcing 'I've been thinking about what we did last time. What I should have done was...' and went on to describe a well controlled experiment. He appeared able to transfer this new found knowledge to new contexts and scored highly in the next two assessments.

At the second and third assessment interviews, we could see a gap widening between the intervention and control pupils, with a far more considered and reflective approaches appearing in the intervention classes. Analysis of the pre-tests demonstrated that both the written and interview assessment instruments were reliable with Cronbach's alpha scores of 0.86 and 0.84 respectively. Scores on the interview based assessments of metacognitive skill correlated highly with the written assessments of metacognitve skill (r=0.67, significant beyond p=0.001). No significant difference was found between the scores of the intervention and control classes in the pre-tests at the 5% level.

Participant observation data collected on the intervention lessons led to three intervention classes and their parallel control classes being invalidated for the purposes of further statistical analysis. In the cases of two classes this was because the teaching approaches used bore little resemblance to those which were intended. In another case, the usual teacher was removed from the intervention class in mid-project to exchange classes with an inexperienced teacher who was having problems. (For further details see Tanner, 1998).

Table 1 shows the mean scores for the active metacognitive skills of planning, monitoring, and evaluating as measured through the written paper and the interview assessment. It can be seen that the mean scores increased for both intervention and control groups over the assessment period. Although both groups exhibited development, the mean of the intervention classes improved by more than the mean of the control classes. Visual inspection of the trend (see Figure 1) indicates the mean of the intervention classes to have accelerated away from the mean of the control classes during the period of experimental teaching, but that after the end of the course the groups improved at roughly parallel rates with the improved performance of the intervention classes being largely maintained in delayed testing five months later.

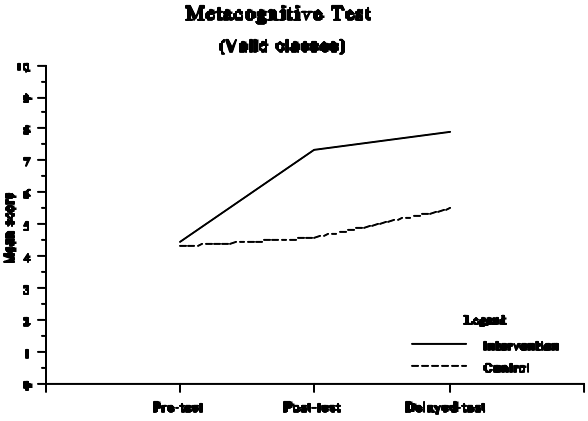
Figure 2 provides the mean scores for the active metacognitive skills of planning, monitoring and evaluating as measured by dynamic assessment interview. The means of both control and intervention groups improved over the period of the assessment (see Table 1). The mean of the intervention pupils had improved by more than that of the control pupils when measured immediately after the course. In delayed testing the intervention pupils continued to exhibit a greater rate of progress than the control pupils. Visual inspection of the trend (see Figure 2) shows the means to be slightly upwards over the period with the difference between intervention and control means widening at both the post-test and delayed-test assessment points. Intervention classes improve more than control classes on average in the post-test and this trend is continued into the delayed test.

Table 1: Means for pre, post and delayed post-tests of metacognition (valid classes)

	Variable	Intervention Mean	Intervention Std Dev	Control Mean	Control Std dev
Written	**Pre**	4.45	3.44	4.32	3.35
assessment	**Post**	7.33	3.86	4.58	3.27
	Delay	7.89	4.15	5.49	3.50
Interview	**Pre**	8.17	5.02	7.22	4.52
Assessment	**Post**	12.20	5.56	8.63	5.77
	Delay	14.19	3.67	9.06	4.71

Multivariate analysis of variance is used taking the pre-test scores as covariates. This adds power to the analysis by adjusting for the small inequalities which existed between groups at the beginning of the quasi-experiment. The effect of type of class is seen in Table 2 which shows that the multivariate F value for the intervention and control classes is significant beyond the 0.1% level for the written test and beyond the 5%

level for the interview assessment with its much smaller sample size. The size of effect is small in both cases, but the intervention was short and as can be seen in Table 3, the effect is sustained long after the end of the intervention.



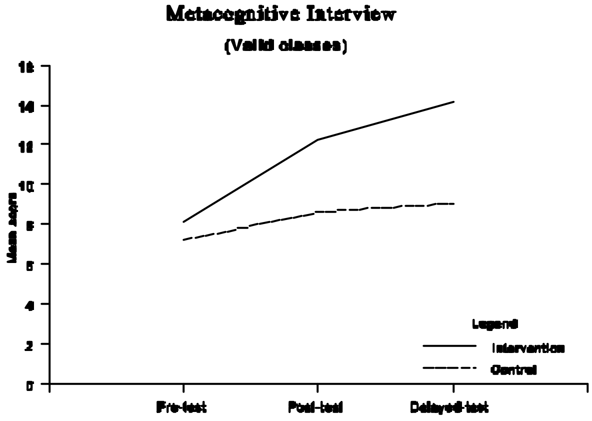


Table 2: Multivariate tests of significance for effect of type of class (Int or Control)

Variable	Hotellings	F value	Error DF	Sig of F	Effect size
Written test	.235	43.67	371	.000	.19
Interview	.374	4.86	26	.016	.27

Table 3: Univariate F tests for the effect of type of class (Intervention or Control)

Variable	Hypo SS	Error SS	Hypo MS	Error MS	F	DF	Sig of F
Written test-2	680.79	3423.26	680.79	9.2	73.98	372	.00
Written test-3	600.43	3989.59	600.43	10.72	55.99	372	.00
Interview 2	90.72	587.18	90.72	21.75	4.17	27	.05
Interview 3	134.95	411.36	134.95	15.24	8.86	27	.01

The adjusted mean scores in table 4 show that although the average scores in assessments of metacognitive skill improved for all pupils as they matured and gained in experience over time the intervention classes gained an advantage over their controls in these key mathematical thinking skills which was sustained well beyond the end of the intervention.

Table 4: Adjusted mean scores for intervention and control groups

Test	Intervention	Control	Significant?
Written test-2	7.35	4.65	√√√
Written test-3	8.17	5.64	√√√
Interview-2	12.50	8.99	√
Interview-3	14.00	9.72	√√

Key: √√√=significant at 0.1%, √√=significant at 1%, √=significant at 5%.

Conclusion

The teachers of the intervention classes were able to teach their classes a number of key metacognitive skills which were sustained long after the initial intervention. Although the size of effect reported here is small, the mathematical thinking skills learned here are significant for real life problem solving and mathematical modelling. Furthermore, there is good reason to believe that when children develop the ability to think in this way, they are able to learn new knowledge more effectively and with deeper understanding (Cobb et al, 1992; Shayer and Adey, 1992; Tanner and Jones, 1995b; 2000). Teaching for thinking is the way forwards.

References:

Brown, A. L. (1987). Metacognition, executive control, self regulation and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65-116). New Jersey: Lawrence Erlbaum. (N)

Brown, A. L., & Ferrara, R. A. (1985). Diagnosing zones of proximal development. In J. V. Wertsch (Ed.), *Culture communication and cognition: Vygotskian perspectives*. Cambridge: Cambridge University Press.

Cobb, P., Wood, T., Yackel, E., & Perlwitz, M. (1992). A follow-up assessment of a second-grade problem centred mathematics project. *Educational Studies in Mathematics*, *23*(5), 483-504.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers*. Baltimore: University Park Press.

Gray, S. S. (1991). Ideas in practice: Metacognition and mathematical problem solving. *Journal of Developmental Education*, *14*, 24-28.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306-355.

Newman, D., Griffin, P., & Cole, M. (1991). *The construction zone: working for cognitive change in school*. Cambridge: Cambridge University Press.

Rowe, H. A. H. (1991). "Observing" thinking and learning processes. In G. Evans (Ed.), *Learning and teaching cognitive skills* (pp. 9-26). Hawthorn, Australia: ACER.

Royer, J. M., Cisero, C. A., & Carlo, M. S. (1993). Techniques and procedures for assessing cognitive skills. *Review of Educational Research, 63*(2), 201-243.

Shayer, M., & Adey, P. (1992). Accelerating the development of formal thinking in middle and high school students III: testing the permanency of effects. *Journal of Research in Science Teaching, 29*(10), 1101-1115.

Tanner, H. (1998). *Using and Applying Mathematics: developing mathematical thinking through practical problem solving and modelling*. Unpublished Ph.D. Thesis, University of Wales Swansea, Swansea.

Tanner, H. F. R., & Jones, S. A. (1993a). Developing metacognition through peer and self assessment. In T. Breiteig, I. Huntley, & G. Kaiser-Messmer (Eds.), *Teaching and learning mathematics in context* (pp. 228-241). London: Ellis Horwood.

Tanner, H. F. R., & Jones, S. A. (1993b). *Hands on maths*. Swansea: University of Wales Swansea.

Tanner, H. F. R., & Jones, S. A. (1994a). The development of metacognitive skills in mathematical modelling. In G. Wain (Ed.), *British Congress on Mathematical Education, 1993: research papers* (pp. 76-80). Leeds: University of Leeds.

Tanner, H. F. R., & Jones, S. A. (1994b). Using peer and self assessment to develop modelling skills with students aged 11 to 16: a socio-constructive view. *Educational Studies in Mathematics*, *27*(4), 413-431.

Tanner, H. F. R., & Jones, S. A. (1995a). *Better thinking, better mathematics*. Swansea: University of Wales Swansea. http://www.swan.ac.uk/education/pgcemaths/btbm/index.html

Tanner, H. F. R., & Jones, S. A. (1995b). Teaching mathematical thinking skills to accelerate cognitive development. *Proceedings of the 19th Psychology of Mathematics Education conference (PME-19), Recife, Brazil, 3*, 121-128.

Tanner, H., & Jones, S. (2000). Scaffolding for success: reflective discourse and the effective teaching of mathematical thinking skills. In T. Rowland & C. Morgan (Eds.), *Research in Mathematics Education Volume 2: Papers of the British Society for Research into Learning Mathematics* (pp. 19-32). London: British Society for Research into Learning Mathematics.

Vygotsky, L. S. (1978). . In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: the development of higher psychological processes*. Cambridge, Mass: Harvard University Press.