Changing Teachers' Classroom Practice through Developmental Assessment: Constraints, Concerns and Unintended Impacts

Rosemary Callingham
University of Tasmania
<Rosemary.Callingham@utas.edu.au>

John Pegg
University of New England
<jpegg@une.edu.au>

Teresa Wright
University of New England
<theresajwright@kooee.com.au>

Developmental assessment processes support the idea of assessment *for* learning and are based on qualitative judgements of students' work. Introducing teachers to a developmental model on which to base these judgements aimed to change teachers' assessment practice. Mathematics teachers in six high schools in New South Wales were provided with professional development over three years that introduced the SOLO model. All teachers reported changes to their practice but these were mainly in the context of teaching rather than assessment. They also identified structural constraints that prevented them from implementing the model.

Introduction

Recent curriculum documents have emphasised the role of assessment in improving learning (for example, Board of Studies, 2002). In particular, increased emphasis on assessment *for* learning, or formative assessment, has become evident (Black & Harrison, 2000). Introducing assessment practices that lead to improvement has been problematic, however, and the promise of improved outcomes from assessment *for* learning has not been achieved on a large scale, despite considerable efforts by education systems (Stiggins, 2007).

One difficulty associated with the notion of formative assessment is that teachers are sometimes unable to articulate their students' understanding with the expectations of the curriculum. This becomes more evident when the process of doing mathematics takes more prominence in the curriculum. Teachers have tended to focus on isolated and easily identifiable skills in preference to higher-order thinking processes (Lai & Griffin, 2001). One possibility for informing teachers is to make use of research into cognitive processes. Despite a large body of available research, however, little use appears to have been made of this evidence in curriculum development or the related assessment processes (Pegg & Panizzon, 2001).

In the study reported here, the aim was to provide tools for teachers to use to increase their identification of the level of development of a particular student, with the ultimate aim of targeting their future teaching more closely to students' needs. Specifically, teachers were introduced to the Structure of the Observed Learning Outcome (SOLO) model (Biggs & Collis, 1982, 1991).

The SOLO model is a neo-Piagetian model that considers the structural complexity of a student's response. It includes two aspects: The nature or mode of the thinking process displayed and the quality, or complexity, of the response within the mode of thinking. There are generally five modes of thinking identified: sensorimotor, which is primarily involved with kinesthetic activity; ikonic which addresses imagery and story-telling; concrete-symbolic which is the mode targeted in schooling; and formal and post-formal thought, which are high level modes typically encountered in experienced adults. Unlike Piaget, the model posits that people continue to draw on modes of thinking typically developed earlier. Within each mode are levels of response, which describe an individual's

increasing sophistication of interaction with the task. Three levels of response can be identified: unistructural (U) where the focus is on a single piece of information, multistructural (M) in which information is strung together sequentially, often in a stepwise process; and relational (R) in which information is processed into a coherent account or generalised within a particular context. There may be more than one U-M-R cycle within a particular mode, and typically two cycles have been identified in the concrete-symbolic mode with respect to a particular concept (Pegg, 2003). The SOLO model is summarized in Figure 1.
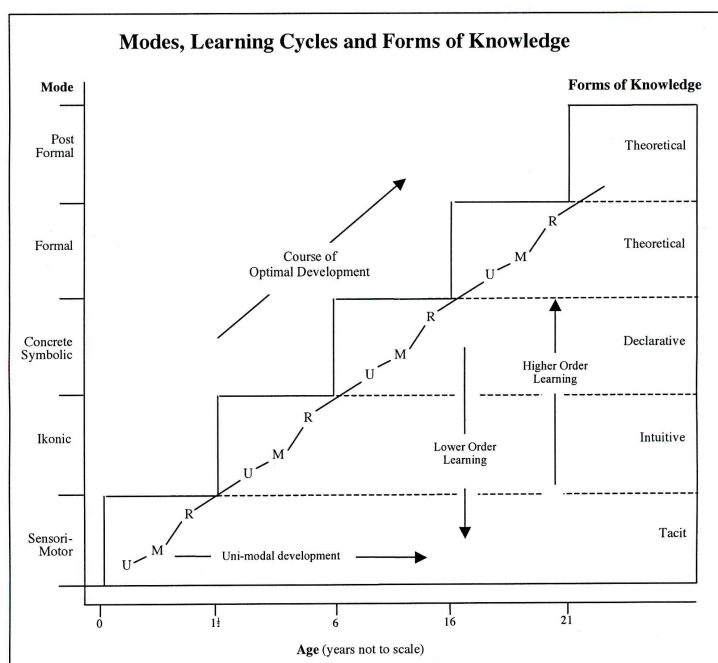

Figure 1. *Modes and levels in the SOLO Model.* (Adapted from Biggs & Collis, 1991)

There have been other studies in which teachers have been introduced to the SOLO model with a focus on characteristics of a U, M or R response (e.g., Panizzon & Pegg, 2008). In the study reported here, however, teachers were introduced immediately to the notion of two cycles, and considerable time was spent identifying shifts from the first to the second cycle within the concrete-symbolic mode. The reason for the decision to introduce two cycles immediately was an attempt to "fine tune" the assessment and make it more sensitive to students' development.

From this background, the research question addressed in this paper is: In what ways do mathematics teachers use a developmental model in their classroom assessment practice?

Method

The study reported was part of a large, 3-year ARC Linkage grant involving teachers of mathematics, science and English. Teachers met twice a year to undertake collaborative professional learning and share experiences, and these events provided one source of data. In addition, a range of documents was collected, including school policies, teachers' planning, examples of students' work and written reflections, and teachers were interviewed formally three times during the period of the study.

The Sample

The findings presented are based on interviews with a total of 11 experienced mathematics teachers in six different high schools in New South Wales. The schools comprised three clusters in different regional areas of the state, and included schools situated in inland and coastal areas, as well as schools within easy reach of large regional centres.

Of the 11 teachers involved, five were Head Teachers of Mathematics, all of whom were male, and six were teachers of mathematics in Years 7 to 10, although most also taught in Years 11 and 12. Of the six classroom mathematics teachers, four were female. The teachers were highly experienced, with all having at least ten years experience of teaching, and most having more than 20 years. They were also generally highly qualified in their subject, most having a major in mathematics in their first degree and a postgraduate Diploma of Education. Two teachers had a Certificate of Teaching, one had a specialist secondary Bachelor of Education, and one was primary trained. All had considerable experience in regional and rural teaching, mainly around New South Wales although two had been in West Australia for a period. Many talked about making a lifestyle choice to teach in regional areas rather than major cities, and appeared to have built their careers around country schools. No teacher expressed a desire to move to Sydney. Overall, the impression gained was of a highly experienced group of teachers committed to rural education.

Data Collection and Analysis

Interviews were conducted three times over the life of the project: at the start to provide baseline data, after about 18 months to provide an indication of how teachers were changing their practice, and at the end to provide a reflective element. Interviews were semi-structured, with a strong framework to provide comparative data across schools and subjects, but with enough flexibility to provide for discussion about issues arising during the project.

A framework for the analysis of the interview data was provided by Niss (1993). He considered assessment from two perspectives: the mode of assessment including all those activities associated with undertaking assessment and reporting on it including the focus of the assessment, the nature of the assessment task, the ways in which judgements are made and the manner of the report. His second perspective was the purposes of assessment which were categorised into three groups: providing information, making decisions and shaping reality. This latter category arose from the privileged nature of mathematics as a "gatekeeper" subject. Exclusion from mathematics could constrain potential life choices. Comments made by teachers in interviews at three different points in time were categorised into modes and purposes of assessment. Changes in teachers' practices and perceptions were identified by comparing these categories from each time point. Constraints on assessment practice were also identified.

Findings

Findings are reported here under the headings of Modes of Assessment, Purposes of Assessment, and Constraints and Concerns.

Modes of Assessment

At the start of the study, most assessment tasks consisted of tests, often given at the end of a unit, as well as half yearly and yearly exams. Assignments or homework were also used for assessment purposes, and one school used a "contract" system where students signed a contract agreeing to complete particular exercises and activities. Some limited mention was made of specialised classroom activities given for the purpose of assessing students' understanding of concepts. It seemed that when asked about assessment, mathematics teachers automatically thought of tests to provide objective information. These tests were almost always teacher developed, although this was not necessarily an individual activity. Many tests were common across a grade, and all teachers who were teaching students in that grade had some input to the test development, although typically a single teacher took ultimate responsibility for writing the test.

The tests usually addressed mathematical content, and many teachers said that the tests they gave had changed little over the years. Multiple choice questions were common, and tests, apart from half-yearly or yearly examinations, centred on some particular aspect of mathematical knowledge such as trigonometry or algebraic equations. No teacher reported attempting to address Working Mathematically outcomes, although they all indicated that the specific outcomes being addressed were made clear to the students and included on the test paper. Other aspects of students' behaviour, such as affective outcomes, were not mentioned at all by any mathematics teacher, although students' attitudes were discussed informally, suggesting that whether the teachers recognised it or not, they did make some judgments about students' behaviour.

By the middle of the project, teachers were reporting more awareness of the need to target assessment more closely to students' development. They were using SOLO terminology to describe what they were doing. For example Stefan commented

> … the R2 was basically when you're working your question backwards – they're given the answers, and they've got to unpack that to come back where – so we've tended to try to have those items now into our tests, or assessments, or class activity – to see if the kids can make that next leap.

Carl also saw the importance of student development indicating "O.K., we're taking a student who's assessed say at U2 level, to R2 level".

By the end of the project, most teachers were making use of open-ended questions, typified by Felix's comment

> I always put an open-ended, thinking type one in if possible, whereas before I didn't offer that – that's a big thing for me! It's improved things no-end. And it helps you set better tests.

Despite the focus on formative processes using SOLO that were emphasised in the professional development, teachers remained focussed on summative tests as the main assessment method. There was some evidence that a few teachers had begun to think more broadly. Stefan commented in his final interview:

> It has made me aware that how we set items, whether it's just a class activity, or it's a mini assessment task, whether it's an activity sheet that we might use towards an assessment task.

The process of making judgements about students on the basis of the assessment information was generally not very clear in the early interview. Marks were awarded in tests and these were generally averaged, sometimes in a weighted form, to provide a final mark. Often some professional judgement came into play, with teachers making adjustments on the basis of specific knowledge about a particular student.

By the middle and end interviews, teachers were more confident about making judgements about students' understanding. In her middle interview, Marion indicated:

> I really look at what the kids have done in their responses. I really have a much clearer picture of where their limitations are.

James also had moved to looking beyond marks by the end of the project:

> I can see that there's more than just a mark for it, in learning – that there's an understanding behind it....I think I can now see questions that hit you, and go: "Bang! That should be marked SOLO-wise."

Reporting processes were strongly influenced by parents' expectations. For example, Felix said:

> … [if] you get a whole heap of writing; the parents don't read it; they just look for A, B, C's, D's, and E's; or places in the class, which they don't get, we don't have any places in class, or marks for juniors. We have A, B, C, D, E, and a comment.

Marion agreed with this perspective stating:

> Well reporting's an issue because I think the expectation is that the parents want more feedback. To be perfectly honest, the reality is the majority of parents are still thinking educationally the way that they were taught.

It appeared that SOLO did have some impact on teachers' approaches to assessment, but possibly less than might have been anticipated. Even where teachers were applying SOLO it was in terms of tests and question types, or in ways of summative marking rather than as a formative approach that could be used to inform teaching. Much of the assessment agenda was driven by reporting demands which, in turn, were linked to real or perceived parent expectations.

Purposes of Assessment

In the early interviews, teachers were not forthcoming about the purposes of assessment. Provision of information, however, was implicit in many of their remarks. Teachers talked about "knowing how kids are going" and being able to discuss this with parents. In most schools, departments did discuss the information provided by the various state-wide numeracy reports, including value-added measures and proportions of students falling into particular bands or categories at school certificate and HSC. There did not appear to be any action taken on the basis of this information, however. Often it was discussed as simply a fact of life rather than as a useful tool for improvement. No head teacher mentioned using assessment information to assist with programming or planning.

By the middle of the project, however, some teachers were thinking more about the way in which using SOLO helped them with planning their teaching. Daniel, for example said:

> But now I think the biggest difference is in my teaching methods before they do their HSC questions. I'm progressing through it and seeing that the kids are at a certain level, and saying to them – and thinking to myself, at school and at home, how can I get them to a higher level?

One teacher who had joined the project late spent considerable time with his colleagues learning what had gone before. In his final interview, this teacher, Patrick, clearly explained how he saw the outcomes of using developmental approaches to assessment:

> … you would actually make an attempt to teach in a slightly different way. If students were at a different level, even though they achieved the same score, then you'd teach them in a slightly different way.

Teachers appeared to have begun to consider formative aspects of assessment, but linked this to teaching practice rather than assessment itself. Marion, for example, described her closer attention to students' responses but discussed this in terms of teaching:

> The main way in which I've changed my teaching, is that I'm now more interested in listening to the responses of my students to these more generalised questions. Rather than 'what's 5 sevens?', I'd say: 'What are 2 numbers that could give me 35?' The product of what 2 numbers will give me 35, or 36, or something…

Decision making following assessment was mainly associated with grading students, and sorting them into different groupings for teaching. All schools operated some form of grading and this appeared to be a source of ongoing discussion among staff. There were diverse philosophies in different learning areas with some faculties opposing grading. In mathematics, however, some form of grading or streaming was universally applied by Year 9, even when classes were heterogeneous in the lower part of the school. Decisions about grading were not taken lightly and in all schools teachers described movements among classes according to merit. Much discussion took place among relevant staff, and, sometimes, parents were also involved in the decision. Little changed over the life of the project. Grading seemed to be an entrenched practice that continued relatively unquestioned by mathematics teachers.

Despite the recognised importance of using assessment for decision making, however, no teacher, or principal, discussed this in terms of Niss's (1993) "shaping reality". The potential effects of decisions made about placement in mathematics on a student's life chances were never explicitly acknowledged.

Constraints and Concerns

Overall teachers were enthusiastic about the project, although they all indicated that thinking in new ways had not been easy at times. As Stefan put it

> I've got to say, for the first year, it really played with my headspace. It was something I couldn't come to grips with. I was really struggling with it, even though I liked it because it was maths.

Teachers were honest about their difficulties, acknowledging that they could not always teach at their optimum level. Lenore said, for example,

> It's infiltrating my classrooms a bit. I find when I get lazy, or I'm under stress or whatever, I just revert back to what's easy for me.

She went on to say:

> In terms of my own teaching, I'm forever frustrated that I don't have enough time and energy into putting into really exploring this more. I find that SOLO and its implications are so wonderful, and the little bits that I have done – I get excited when I come out of the classroom, and I see what has happened. I would love, particularly if I was to go back to starting teaching all over again, I would love to have brought this into the classroom right from the very beginning. It's like, unfortunately, teaching an old dog new tricks.

Time was, not surprisingly, often mentioned as a constraint by all teachers. This was not, however, the only issue. Many teachers commented on the pressure from high level examinations at the end of schooling. Repeatedly they indicated that the use of SOLO helped them teach top level students better, typified by these comments:

> It changed my approach to 4-unit Maths teaching. From then on the curriculum became far less important, and it was: "I have to teach them how to think." I've never really looked at it like that. And I know I'm right. (Felix)

> I reckon it's benefited the kids 100%. I reckon if I was loaded up with some kids like I did this year…and next year I've got the 11 Extension again…I reckon it just benefits them, because I'm just thinking about pushing them through the levels all the time. (Daniel)

Notably, however, teachers did not see SOLO as useful for students at the lower end of the ability spectrum. This comment by George exemplified many teachers' thinking:

> I think it's really good for smart kids. I'm not sure it's very useful for my bottom group of kids.

Some teachers didn't believe that their practice had changed greatly across the years of the project. They did, however, acknowledge that being introduced to SOLO as a developmental model had focussed them. A typical comment was this one from Richard:

> It hasn't had a major impact. It's probably refocused me, and idea that I think I've always had, but didn't have in these terms.

Despite the difficulties of time, and the constraints of examination requirements, teachers worked hard to implement new ideas in their assessment and teaching. Their efforts were rewarded when they found that their students achieved better and demonstrated deep understanding of a topic. James, for example, became very excited when he described a test set for Year 10 students:

> Like we gave a test in Year 10 – we gave an assessment test in Year 10, and we gave it on 2 cylinders. All we did was make it a really long thin cylinder, and a short fat one. The kids had to write about cylinders; had to write down what they understand about their volume and capacity, and what impact, in lowering the height and expanding the radius had on their capacity. And that's better than a kid just learning $\pi r^2 h$. Some of the responses were incredible. Kids went right into it, and went into ratios, went into r^2 …

Stories like this, shared at project meetings, provided impetus for teachers to continue with new approaches to teaching, even when other colleagues in their schools were not involved.

Discussion and Implications

A number of issues emerge from this study. Initially the project was intended to change assessment practice with a particular focus on understanding students' responses in order to target teaching and feedback more effectively. Over the period of the study, however, the teachers involved caused the centre of attention to move more to teaching and classroom practice than assessment itself. Much of their practice, particularly in the questions asked, could have been described as assessment *for* learning but it was evident that the teachers themselves regarded teaching differently from assessment. When pushed to talk about how the use of the SOLO developmental model impacted on their assessment practice they spoke only of item setting for tests, ignoring the comments they had made about the nature of the questions they asked students in their teaching. This finding suggests that the push to improve outcomes through formative assessment might have been more effective if it had been placed within a teaching rather than an assessment framework. Assessment was driven by external pressures, such as high stakes examinations and reporting requirements, and appeared resistant to change. In contrast, teachers were in control of their classroom practice, and were willing to try out new ideas and activities.

Few teachers used SOLO with lower ability classes. They saw it as a tool for extending high achieving students or not limiting these students to lower level responses. This unexpected finding may have been an unintended consequence of a focus on identifying the second UMR cycle in concrete-symbolic mode, or on moving students into formal mode in the higher grades of schooling. Because teachers were introduced to the two-cycle

model initially, identifying subtle differences between a first- and second-cycle multistructural response, for example, absorbed much discussion time during project meetings. Teachers, however, appreciated the additional information provided by identifying a second cycle, and the rich discussions during the meetings were often referred to by teachers in their reflections on the project. To implement developmental perspectives, whether in teaching or assessment, across the ability range, however, may require a more deliberate focus on lower levels of response from students.

The impact of external pressures such as parental expectations and examinations was strong. Unless such factors are reduced or changed it seems unlikely that teachers will change their assessment and reporting practices, regardless of policies that aim to promote change. Within these constraints, however, it was clear from the teachers' responses that the developmental framework provided by the SOLO model was valuable and relevant.

Acknowledgement

References

Biggs, J. & Collis, K. (1982). *Evaluating the Quality of Learning: the SOLO Taxonomy*. New York: Academic Press.

Biggs, J. & Collis, K. (1991). Multimodal learning and the quality of intelligent behaviour. In H. Rowe (Ed.), *Intelligence, Reconceptualization and Measurement* (pp. 57-76). New Jersey: Laurence Erlbaum Assoc.

Black, P. & Harrison, C. (2000). Formative assessment. In M. Monk & J. Osborne (Eds.), *Good practice in science teaching: What research has to say* (pp. 25-40). Buckingham, MK: Open University Press.

Board of Studies (2002). *Mathematics years 7–10 syllabus advice on programming and assessment*. Sydney: Author.

Lai K. & Griffin, P. (2001). *Linking cognitive psychology and item response models*. Paper presented at the Australian Association for Research in Education Annual Conference, Brisbane, Queensland. Retrieved 6th February, 2007 from http://www.aare.edu.au/01pap/lai01487.htm

Niss, M. (1993). Assessment in mathematics education and its effects: An introduction. In M. Niss (Ed.), *Investigations into assessment in mathematics education,* (pp. 1-30). Dorderecht, Netherlands: Kluwer Academic Publishers.

Panizzon, D. & Pegg, J. (2008). Assessment practices: Empowering mathematics and science in rural secondary schools to enhance student learning. *International Journal of Science and Mathematics Education, 6*, 417-436.

Pegg, J. & Panizzon, D. (2001). Determining levels of development in outcomes-based education: Nice idea, but where is the research-base for the decisions taken? Paper presented at the American Educational Research Association Conference in Seattle, on 10-14th April, 1-5.

Pegg, J. (2003). Assessment in mathematics: A developmental approach. In M. Royer (Ed.), *Mathematical Cognition* (pp. 227-259). Greenwich, Connecticut: Information Age Publishing.

Stiggins, R. (2007). Assessment through the student's eyes. *Educational Leadership, 64*(8), 22-26.