PAN07110

Shifting sands: Using SOLO to promote assessment for learning with secondary mathematics and science teachers

Debra Panizzon[1], Rosemary Callingham[2], Terry Wright[3], & John Pegg[3]

[1]Flinders Centre for Science Education in the 21st Century
Flinders University
Adelaide, South Australia
debra.panizzon@flinders.edu.au

[2]School of Education, University of New England
Armidale, NSW

[3]National Centre of Science, Information and Communication Technology, and Mathematics Education for Rural and Regional Australia (SiMERR)
University of New England
Armidale, NSW

ABSTRACT

Assessment *for* learning places greater emphasis on assessment being linked intrinsically to the teaching and learning process. In this context summative and formative assessment become critical in providing feedback to teachers and students about *what* is understood and *where* learning and subsequently teaching should be directed to enhance student understanding. In NSW, this focus has required a major shift in teacher thinking and practice. A study using the Structure of the Observed Learning Outcome (SOLO) model as a theoretical framework was developed to provide twenty-three secondary mathematics and science teachers with a structure and language to reconsider their assessment and teaching practices for Years 7-10 during a professional development programme. This paper describes the changes in teachers' practices by comparing baseline and mid-project interviews. Overall, mathematics teachers applied the model to identify developmental pathways in mathematical conceptual understanding as a means of improving their teaching strategies in the classroom. Alternatively, science teachers recognised the usefulness of the SOLO model for planning a range of assessment tasks that explored students' scientific understandings and then implemented this knowledge to restructure their science programmes. These results demonstrated a major shift in teachers' views of assessment and their practices from the initial interviews conducted at the outset of the study.

INTRODUCTION

Assessment is a central part of teaching (Linn, 1990) with the majority of teachers using it to gauge students' knowledge, understanding, and skills at a particular time in a learning sequence (Stiggins, 2002). In the disciplines of mathematics and science in secondary schools this is achieved most often using topic tests, examinations, or research assignments (Goodrum, Hackling & Rennie, 2001). However, application of these traditional assessment tasks tends to:

- reduce a discipline (e.g., science) to learning about facts;
- lower the level of cognition required by the student;
- inhibit student questioning as they maintain focus on the topic being taught;
- ensure that most teachers 'teach to the test';
- restrict teaching to aspects that are assessable in this manner; and
- ensure that creative and innovative teaching methods are omitted as teachers 'violate their own standards of good teaching' (Black, 1993:52).

Changes to these traditional assessment methods have been a world-wide focus over the last decade (Morgan & Watson, 2002) with the work by Black and Wiliam (1998) introducing the notion of assessment *for* learning in contrast to the prevailing assessment *of* learning perspective. In their view, assessment must move away from the summative regime identified above, based around the collection of marks for accountability and reporting purposes to one that is integrated into the teaching and learning process. Biggs (1996) referred to this idea as 'constructive alignment' with curriculum, pedagogy and assessment linked to ensure that each component is used to inform the direction of teaching and so enhance student learning. A critical element in this approach is the importance of ongoing or formative assessment to monitor on a day-to-day basis where students are and where their learning needs to be directed (Bell & Cowie, 2001; Hattie & Timperley, 2007; Pelegrino, Chudowsky, & Glaser, 2001; Shepard, 2000).

Regardless of whether assessment is used summatively or formatively, teachers are ultimately responsible for making judgements about the quality of students' understandings (Linn, 1990). While Morgan and Watson (2002:80) considered that 'teachers are in the best position to ensure that assessment is equitable', they recognised that teachers require experience and training to ensure a high degree of consistency in their judgement. Koetz (1998) supported this view suggesting that rubrics provided one means of providing teachers with a set of well-developed criteria against which they could assess students' work. However, Wilson and Sloane (2000) recognised that a critical component to the success of rubrics or any other process used to gauge student understanding is the need for a theoretical framework that provides a development perspective of student learning. In their view utilisation of such a framework helps to move assessment away from one-shot tests to 'an approach that focuses on the process of learning and on an individual's progress through that process' (Wilson & Sloane, 2000: 183).

This paper addresses the particular aspect of focussing assessment on the process of learning, presenting results from a study in which secondary mathematics, science and English teachers were introduced to a developmental theoretical framework to support changes in their assessment practices to better meet the needs of their students. Given that the study is in its third year (with one year to go), we present the findings from mid-project interviews with mathematics and science teachers and compare them to the initial views expressed by the teachers at the outset of the study. Two case studies representing longitudinal changes and

impacts on the practices of two teachers (one science and one mathematics) over the two-year period are provided. The results highlight the shifts teachers made in reconceptualizing the purpose of assessment, the types of tasks generated to assess their students, and how they used results from these tasks to alter their pedagogical practices.

THEORETICAL FRAMEWORK

The Structure of the Observed Learning Outcome (SOLO) (Biggs & Collis, 1982, 1991) model has much in common with the neo-Piagetian frameworks of Case (1992), and Fischer and Knight (1990). SOLO was developed to focus on the structure of students' responses after a learning experience. Underpinning the model is the assumption that cognitive understanding does not equate to a stable cognitive construct as suggested by Piaget (1954), but involves individual characteristics that are content and context dependent. The variables affecting the quality of a response include the availability of working memory, the amount of information that can be retained by the learner, and specific features of the learning task (Biggs & Collis, 1991).

The SOLO model embraces two important features. The first concerns the nature or abstractness of the response and is referred to as the *mode of thinking*. This refers to the type of intellectual functioning that is required to address a particular stimulus. As such, each mode has its own identity and specific idiosyncratic character. The second feature depends on an individual's ability to handle, with increased sophistication, relevant cues. This feature is referred to as *levels of response*, which are seen to reside within cycles of learning that provide a hierarchical description of the nature of the structure of a response. While these levels occur within each mode, the specific nature of these levels is dependent on the particular mode targeted by the stimulus item.

In terms of the *modes of thinking*, five are identifiable (Biggs & Collis, 1982).

Sensorimotor — This mode becomes available at birth. It encompasses the coordination of actions and the learning of motor skills within the physical environment. As this type of knowledge involves 'knowing how' to complete a physical task, it is termed tacit knowledge. Importantly, this mode plays a role throughout life particularly in relation to the development of sporting or other physical skills.

Ikonic — Accessibility to this mode occurs as actions become internalised by the individual resulting in the use and development of language and imagery. Young children in this mode use stories and mythical characters to explain human interactions while adults use the mode to assist in appreciating music and art. Thinking intuitively is a good example of adults working within the ikonic mode.

Concrete symbolic — Working in this mode involves a major shift in abstraction as concepts and operations are applied through the symbolic systems of written language, number, and musical notation. These systems involve an internal logic and order or 'knowing what' as individuals describe their experienced world using symbolic descriptions. It is this mode that is addressed most commonly with learning in the upper primary and secondary school.

Formal	Within this mode, individuals seek to understand the relationships between concepts as their thought processes become more abstract and they move away from the need for concrete referents. They are able to question ideas and formulate hypotheses within specific disciplines. This is considered as the highest level of abstraction required in professional practice and at an undergraduate level.
Post Formal	This mode is demonstrated in postgraduate study where the conventional principles of a discipline are questioned and explored further so that the knowledge of the discipline expands. There is some debate, however, about the existence of this mode.

While the five modes of thinking are distinct and are seen to become available in the order provided above, the functioning in a later acquired mode (e.g., concrete symbolic) does not preclude the use of an earlier acquired mode to support understanding (such as ikonic or sensorimotor). This is referred to as multi-modal functioning. The important consideration here is that all modes are available and continue to develop throughout life in response to experiential, social, cultural, educational, or genetic factors (Collis, Jones, Sprod, Watson, & Fraser, 1998; Pegg, 2003).

In terms of the second feature of the model, three *levels of responses* make up a cycle of learning within a mode. These include:

Unistructural (U)	Indicates that the individual has understood the task but can relate only one piece of relevant information. While the content of the response may be consistent with the data, inconsistencies often emerge within the response itself.
Multistructural (M)	Occurs when the individual is able to identify two or more pieces of data relevant to the question but there is a lack of integration.
Relational (R)	Demonstrates that the learner is able to identify a number of elements that are consistent with the question and relate them to one another around a particular concept. However, this level of response does not take general principles into account so that generalisations are not consistent within different contexts.

The modes of thinking and levels of responses within these modes are summarised in Figure 1.

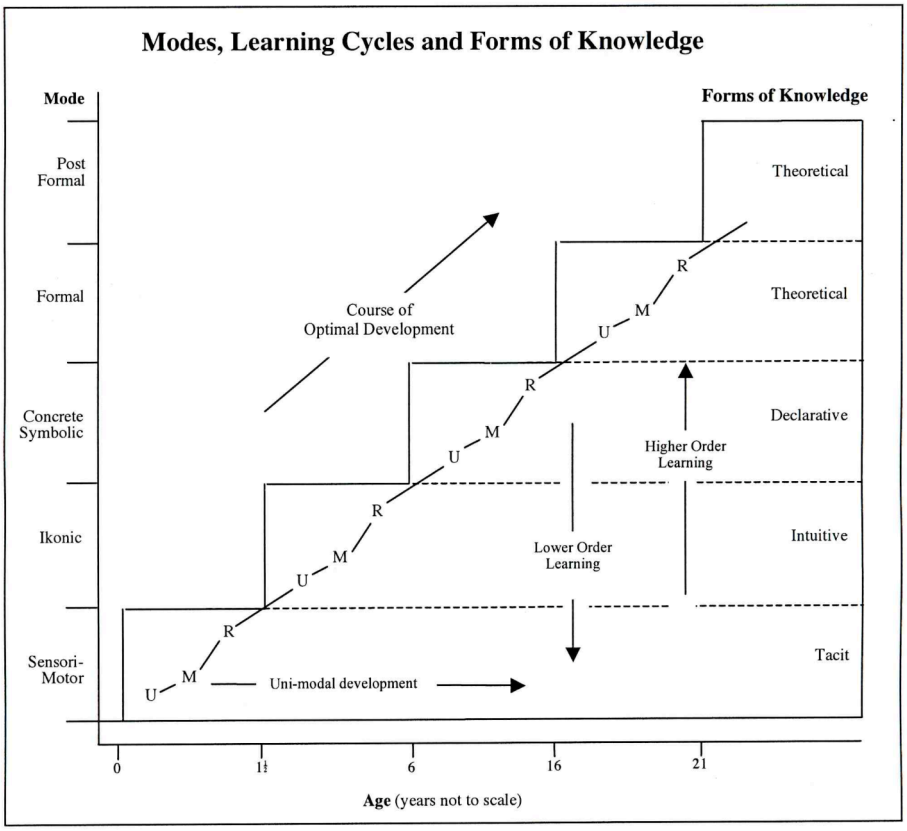

Figure 1. Modes and levels in SOLO Model.
(Adapted from Biggs & Collis, 1991)

The U, M, and R levels identify a cycle of understanding within any particular mode. Research conducted over the last decade (Collis *et al*. 1998; Panizzon, 2003; Pegg & Davey 1998; Pegg, 1997; Watson, Collis, Callingham, & Moritz, 1995) indicates that a single learning cycle (i.e. U-M-R) is not sufficient to explain fully the development of concepts or the diversity evident in responses within a mode. Subsequently, an additional learning cycle of development or understanding emerged in the concrete symbolic and formal modes. As a consequence of this research, two unistructural-multistructural-relational cycles (i.e., U_1-M_1-R_1, U_2-M_2-R_2) have been incorporated into the model for the concrete symbolic and formal modes (Figure 2).

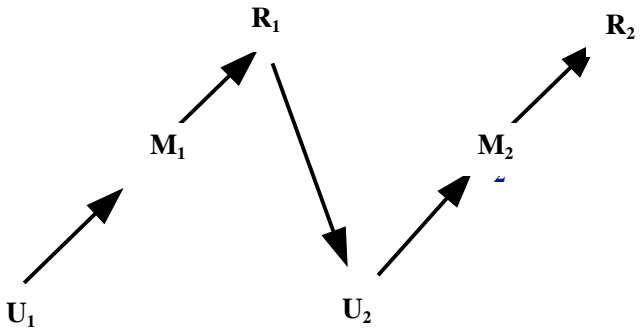

Figure 2. Two learning cycles in the concrete symbolic mode.

The SOLO model is relevant to the present study because it provides a framework to describe the underlying structure of understanding demonstrated in students' responses. By utilising this framework, teachers are potentially in a position: (i) to better target their teaching strategies to the 'level' of the students; and (ii) to critically analyse the faculty teaching and assessment programme from a developmental perspective.

CONTEXT FOR THE STUDY

Within New South Wales (NSW) the Board of Studies develops the syllabuses used in all schools from Kindergarten to Year 12. It also administers two major public examinations in secondary schooling: (i) the Higher School Certificate (HSC) at the completion of Year 12; and (ii) the School Certificate (SC) at the end of Year 10. Over the last seven years in NSW major changes have been made to the syllabuses for Years 7-12 with a focus around *assessment for learning* (Black & Wiliam, 1998). Driving this agenda is the need to 'enhance teaching and improve learning' (Board of Studies, 2003: 5) by assessing subject outcomes using a variety of assessment tasks while providing detailed, meaningful feedback to students.

Discussions with teachers across NSW have highlighted that many are confused by an apparent conflict between the 'high-stakes' assessment agenda driving the assessment practices and accountability in the HSC and the *assessment for learning* agenda being implemented in Years 7-10. In some schools, many of the processes used in Years 11 and 12 are being introduced into the junior years as a means of ensuring that assessment is fair and equitable.

This study was devised to support teachers and minimise this apparent conflict using the SOLO model to guide assessment and teaching practices, particularly in Years 7-10. Over a three-year period mathematics, science and English teachers were provided with ongoing professional development around assessment and the SOLO model. This allowed them to share their experiences, strategies and ideas with teachers in their own schools and across school districts in NSW. In this paper we focus on mathematics and science teachers only.

METHOD

The following Research Questions guided this particular phase of the study.

1. What changes in assessment practices are identifiable for mathematics and science teachers after two years professional development?
2. How does a working knowledge of SOLO as a theoretical model impact teacher assessment practices?

Research Sample

The sample for the study included two teachers from each of the two discipline areas from six schools in NSW. The schools were selected to ensure there were two schools in each of three districts to allow local meetings of teachers to occur. Ultimately, this formed three clusters of two schools representative of rural, regional, and coastal areas of NSW. A summary of information about the teachers is provided in Table 1.

Inclusion of Head Teachers was considered critical to the study given the important role they play in terms of leadership of the faculty in NSW schools. To initiate change and ensure the sustainability of any changes made within the faculty it was paramount that they be part of the team. All of the teachers were highly experienced practitioners with many having taught in a

number of different schools during their careers. Overall, they represented a group of committed teachers who were keen to adopt 'new ideas' and challenges as a means of improving student learning and engagement.

Table 1.
Demographic information about participants

Discipline Area	Position in School		Sex	
	Head Teachers	Classroom Teachers	F	M
Mathematics	5	6	4	7
Science	6	6	4	8
Total	11	12	8	15

Design of Study
The project was funded by the Australian Research Council (with the NSW Department of Education and Training as an industry partner) for a four-year period, with three of these years involving professional development. Once schools were identified and teachers had agreed to participate in the project, each school was visited by the research team. During this visit, each teacher and principal was interviewed to explore aspects of assessment practices and issues for teachers. In addition, copies of assessment tasks and assessment policy documents were collected from each discipline grouping of teachers. After this teachers engaged in ongoing professional development provided by the research team (Table 2).

Table 2.
Timeline for the project over the two-year period

Month	Instructional and Research Activity
2005	
March	Introductory visit to schools by the research team. Conduct base-line interviews with teachers and principals in each school.
August	Professional development day in each cluster with attendance by all teachers, some principals and representatives of the NSW DET (the project industry partner). Introduction to the SOLO model and workshops about issues related to assessment. Teachers left the workshop with an activity to complete over the next few months.
October	Follow up to professional development day - Progress Report on tasks.
December	Summit held at the local university bringing together all participants in the one location over a two-day period. Presentations by the research team, interactive workshops around the SOLO model and development of assessment tasks. Planning for 2006 with teachers in each of the discipline areas deciding on an activity to undertake in next year.

2006	
February	Feedback provided by participants on project activities and their 'journey' during 2005.
March	Contact visits to all schools by a representative of research team. The Head Teachers and principal in each school were interviewed about the impact of project activities in school.
May	Professional development day in each cluster with attendance by all teachers. Follow-up to the SOLO model and workshops about assessment. Teachers in each discipline group targeted an area of focus for the next six months.
August	School visits by members of the research team to conduct mid-project interviews. All participants were interviewed individually to obtain a snapshot about their understandings of the SOLO model and any internalisation that might have occurred as a result of the activities so far.
December	Summit held at the local university bringing together all participants in the one location over a three-day period. Interactive workshops in various discipline groups around the SOLO model, development of assessment tasks, and report-back from teachers about changes to assessment in their individual schools. Planning for 2007 with teachers in each of the discipline areas deciding on an activity to undertake in final year.
2007	
February	Feedback provided by participants on project activities and their 'journey' during 2006.

In addition to the design outlined above, ongoing consultative help was available for the teachers between workshop sessions during both years. This support was provided by our project officer who maintained communication with teachers via email, phone and fax. Consequently, every attempt was made to meet the needs of teachers within the constraints of the budget including some financial support to cover teacher relief so that teams of teachers within each school could meet and work through a particular activity.

Data Collection and Analysis

Four methods were used to collect data for this study:

- Interviews were conducted with each teacher and principal during an introductory meeting to each school. These semi-structured interviews helped to establish a rapport with the teachers and provided baseline data around teachers' current assessment practices. Mid-project interviews were completed eighteen months after the initial interviews by the research team.
- Professional development workshop discussions were recorded.
- Documentation produced by teachers in each school and handed out to other teachers was collected.
- Surveys were completed by each teacher at the end of the yearly Summits in 2005 and 2006. These required teachers to reflect on the changes and developments that had occurred in their own assessment and teaching practices during the year.

In terms of analysis, all of the interviews were recorded and transcribed. After the research team read a number of the transcripts a series of major concepts emerged (Glaser & Strauss, 1967; Strauss & Corbin 1998). The team developed nodes and sub-nodes using NVivo assisted by the project officer. Follow-up discussions over a number of meetings ensured consistency of coding and helped to validate the nodes being developed. Additionally, notes were made regarding each of the professional development workshops with sections that were directly relevant to assessment and teaching practices being transcribed. The teacher surveys and documentation were analysed using content analysis to identify themes similar to and different from the interviews. Data from all sources were triangulated to develop the findings presented below. Excerpts quoted are verbatim with pseudonyms used where appropriate to ensure anonymity of the participants and schools.

FINDINGS

Coding and an analysis of the data identified several themes in relation to assessment and teaching practices. In this section we present an overview of the baseline data for mathematics and science teachers and then a summary of the mid-project interviews for both groups of teachers. Two case studies for Barry (mathematics teacher) and Charles (science teacher) are presented to highlight in a longitudinal manner how these teachers altered their classroom practices over the two-year period.

Mathematics Interview Results

Baseline data

At the start of the project, the majority of mathematics teachers perceived assessment in terms of summative tests. They indicated that assessment had changed little over the years despite the various initiatives introduced, such as Quality Teaching (Professional Support & Curriculum Directorate, 2003) and an outcomes-influenced syllabus. Reporting to parents appeared to be the biggest driver of assessment, with teachers in all schools describing reporting processes in terms of comparisons with other students, either within the class or across the grade. Mathematics teachers, in general, were perceived as the slowest to change. One principal commented about mathematics assessment 'In the mathematics area probably there's been no change at all in my time – or even before my time' (CMPRIN). Mathematics teachers, however, saw the issue rather differently. They saw themselves as experts and did not tolerate interference:

> I think there's a reluctance on mathematics teachers' behalf to be dictated to about how they will deliver their curriculum. They like to know, 'this is what we want the kids to know in the end, and I'll be the best judge of that' (CMMT2).

Mathematics teachers themselves conceded that they were conventional and conservative, but also very efficient:

> Like all mathematics staffs we're very process oriented; very process driven. You know – grind marks; churn them out, grades; that sort of stuff; get everything done; very efficient – all of them. And this is the reality (CDMHT).

The notion of developmental assessment for many of these teachers was unusual. They were used to providing a mark, often using very sophisticated schemes based on the teachers'

estimates of the difficulty of the items, and some form of part-marks if the question was partially correct. These marks were sometimes scaled using means and standard deviations. There was little collaborative marking; rather there was a tacit assumption that all teachers applied the marking scheme consistently. The meaning of the marks was not linked to any notion of student development, although some schools did deliberately target outcomes.

Assessment was driven largely by Higher School Certificate requirements. Where the HSC expectations changed, such as in the nature of the questions in the examination or in the tasks set, teachers changed their practice accordingly, and there was a 'trickle down' effect (Stephens, Clarke, & Pavlou, 1994) to lower grades.

In the classroom, few teachers described any conscious formative assessment. One teacher, who did attempt to formalise classroom assessment, recognised the difference between formal and informal assessment but thought that in general it was undertaken somewhat haphazardly:

> I suppose your informal assessment is far different, and it varies, depending on the knowledge of the particular staff member. I mean some staff members probably don't do a huge amount of informal assessment, or they do it without thinking about it too much (CDMT1).

Another teacher described classroom assessment, although informally conducted as a summative process:

> I mean, all I can say is, that in the classroom, I'm wandering around, and I see kids doing things, and there're kids that I need to help, and there's kids that I don't need to help. I mean, as far as I'm concerned, I'm getting a good idea of where the kids are achieving their outcomes in that respect (CMMT1).

In summary, mathematics teachers at the start of the project were perceived in general as the most conservative faculty by both principals and themselves. Assessment was linked strongly to external examination requirements and predominantly summative in nature. There was a strong desire to do the best by their students, and a feeling that many new initiatives had been and gone over the years, creating a reluctance to change.

> And then you've got the situation where you're going through and you're using all these different terms, where everyone has to be retaught – metalanguage, and all the rest of it. Quite frankly, they don't want to do it (CDMHT).

Mid-project data

Mid-project data indicated a considerable shift in mathematics teachers' thinking. They were using SOLO terminology with confidence and applying SOLO ideas in their teaching and assessment, as shown in the following extracts from interviews with teachers in three different schools.

> It's made me think more, when kids come up with an answer, whether they really understand the answer they're getting, and that, and I think that's where SOLO helps (CLMHT).

> I think SOLO has made me open my eyes, and we're especially looking at the moment at the new reporting system, where we're supposed to go A to E, and if you look at an A, it has to be something like 'Outstanding', and I really don't think, with our mode of testing, we can actually say a student IS outstanding, unless we put some sort of SOLO activity in a test, where we can say: 'Yes, that child IS outstanding! He understands not only the skills that I've taught him, he's understood the whole concept behind whatever it happens to be' (CFMT1).

> Once you've established where they're at, we've got to move these kids from unistructural to multistructural as quickly as we can, so go back to the drawing board; work with that group; let the rest of the kids feel like they're being rewarded, because they have got it (CVMT1).

Several teachers indicated that learning about SOLO had caused them to rethink how they were teaching:

> It changed my approach to 4-unit mathematics 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. I'm getting more successful at being a 4-Unit mathematics teacher now. The curriculum's just a means to get them to think analytically (CHMT2).

> I now think I really look at what the kids have done in their responses. I really have a much clearer picture of where their limitations are. I think I'm more looking at where their limitations are, rather from a SOLO – well that sort of fits into the SOLO thing, in the sense I'm seeing 'Well – this child really needs a lot more work in this area', because they are not – this is where they are stuck, and they are just not moving on (CVMT1).

These comments suggested a significant shift in mathematics teachers' notions of teaching and assessment. No longer were they driven only by curriculum demands. Instead they were recognising students' needs and attempting to meet these needs in different ways appropriate to their situations.

Mathematics Case Study: Barry (teacher)

The changes in mathematics teachers are exemplified by Barry. Barry is a very experienced mathematics teacher who has been at Crystal Mountain High School for 22 years. Over his career he has taught several subjects, including computing. He holds a mathematics major in his degree, gained through study at a Teachers College, with additional qualifications in computing, obtained when computers were introduced into schools.

During the baseline interviews, Barry recognised his conservative approach to teaching mathematics:

> I have pretty entrenched feelings about how to deliver the curriculum to the kids; how I deal with it. I suppose after 30 years you're a bit like that – you can't help that.

This conservatism was supported by his view of mathematics and mathematics teachers generally:

> Maths hasn't changed. I mean algebra is algebra.
>
> I guess maths has been reasonably – out of all the faculty areas, I think maths has probably maintained the status quo more than anywhere else.
>
> [There are] difficulties adapting to our area, and I guess maths teachers are a little bit more rigid in how they see the world in schools.

Teaching in computing as well as mathematics, Barry saw inherent differences between the two subjects, and the way in which he approached them:

> I think I'm a different person, from one room to the other. I don't try to be, but it's just the nature of it. My maths class is very – I guess rigidly controlled – I think I'm probably one of the two people in the school that the kids think have got the most control in the classroom. My maths classroom is very quiet, and very controlled. Whereas my IT room – I'm quite happy for there to be noise there, because that's better for productivity – but I don't think it is in a maths classroom.

Although Barry did not use the term formative assessment, he did try to ensure that his students understood the assessment process and devised ways in which they could learn how to study for tests. He described in considerable detail a shift he had made in his assessment practice following a test in which his students had done very badly. He gave the students the test one week before, corrected it and returned it, and deliberately pointed out to the students where they had made mistakes. After one week he gave the same test but with some restrictions:

> I said, 'The only rule is, you can't bring the written version you've done at home in with you. You've got to come in and do it there and then.' Oh, the other thing I said was 'The pay-off is, for every question you get wrong, if it was worth 2 marks, I'll deduct 4.'

His rationale for this approach was to help his students study for formal, external assessments:

> I was trying to get them into that how to practice and to study for an exam. It seemed to work. The group of kids – when it came to major assessment tasks I couldn't do it that way of course, but it got them into the mode of how to study for tests,

Barry saw assessment in mathematics as traditional and driven by reporting as well as examination demands. When asked about the information provided on reports, he agreed that it included marks, percents and rank saying:

> Yes. It's all straightforward, traditional maths assessment, that one. There's a reporting issue there too. Parents like to know how their kid's going amongst their cohort – whether they're the top or the bottom.

He described mathematics teachers as attempting to fit their assessment to new initiatives, rather than changing their practice:

> You know how I said earlier that mathematics teachers generally don't like to be told how to teach this; how to deliver this? It's that reluctance. And no matter what the assessment system is, it's 'How can we adapt what we are doing to fit in with that?' rather than, 'Do we need to go back to square one, and start again and say 'Do we need to change what we deliver and how we deliver it?' to be able to meet that?' It's just 'can we fit that in? Can we peg that in with this?' And if we can we do. That's how I think it is in that area.

He changed his teaching practice to address the different needs of his classes. With lower ability students, recognising their lower motivation, he aimed to get them to enjoy mathematics lessons in order to promote learning.

> I guess it's different grades, different levels of kids. You have a little bit of a different focus. But with those kids, they came not interested; third class down out of 5, and I think if they can come to maths and be happy about it, and the classroom's a comfortable place to be, then you've got a chance for learning to happen.

Hence, at the start of the project, Barry was a highly experienced, and successful, mathematics teacher. He could articulate clearly differences between mathematics and other subjects, and rationalised these variations in terms of subject requirements and wanting to do the right thing by his students.

In the mid-project interviews, however, there was a substantial shift in Barry's thinking. He had found learning about developmental approaches to be very positive:

> So SOLO's been really interesting to me for that – looking at the cognitive processes. I didn't like that when I was at college, but I've really found that interesting and affirming.

Although the project was focussed largely on formative assessment, Barry saw SOLO as being linked inextricably to teaching:

> To me it's more about looking at the whole learning process, not just the assessment part of it, and certainly, writing assessment tasks that focussed on that.

He retained his focus on students and still recognised the importance of ensuring that they had the appropriate skills to address the assessment task:

> I think it's about how you deliver the material and – because, even if you were just do the assessment tasks, you'd have to – not train the kids, but give them that exposure to it in the classroom, to that type of task, so they're aware of how they should be thinking, and what sort of answers they should be responding with. Because it's different, so for them, in terms of how they respond to not only written things, but verbal things.

Barry recognised his changed approach to teaching and described his prior teaching practice in these terms:

> You'd set your path and once you'd prepared all your lessons, you just went with it, even if they weren't working that well, and the kids weren't learning it that well.
>
> With my trigonometry thing, I did it once; and then did it again and as soon as I did it I went, 'No, no, that probably needs' – it's not a just write-it-once thing. I think you do it, and then you look at it later, and you think: 'Gee, if I had just done that, that would have made it even better!' I think it's a real evolutionary thing, because of the nature of the task. Because you're trying to lead them – not lead them, but you're trying to give them an avenue of showing, 'Am I here? Am I here? Am I here?', so you want that path to be fairly clear for those groups, I reckon. So that's how I think about it.

Given Barry's very conservative approach at the beginning of the project, and his reluctance to make changes, these comments indicate a significant transformation in a highly experienced teacher towards the end of his career. He has become more reflective and responsive to new approaches, but has not lost his focus on the needs of his students.

Science Interview results

Baseline data

Most science faculties in schools had adopted assessment practices in years 7-10 that resembled increasingly those employed for years 11 and 12. These practices emphasised common written tests, common practical tests and formal half-yearly and yearly tests or examinations across a cohort of students. One of the most strongly identified issues among the science teachers was an emphasis on accountability and justification of assessment procedures and reporting. For example, to ensure that students completed tasks on their own, all assessment tasks in the school were undertaken during lesson time:

> We have a faculty policy that we don't do any at-home assessment. It's all got to be done in class. From strictly all in-class towards the end of the school, where they can do their own research at home. They have to bring in their rough notes etc. A couple of reasons behind that. Whose work are we assessing? So we make sure we're assessing kids' work (CLST1).

Concerns about accountability and the need to implement objective measures that could be used to account for descriptions of student performance were common. Strategies employed to enhance objectivity and consistency included having an across-the-cohort task set by only one person, often with input from colleagues during the design phase. In some cases all marking of a given task was being conducted by one person, but generally marking was delegated to the class teacher, who marked his or her own class in accordance with a marking scheme that was either centrally-formulated or developed collaboratively. There was little evidence of teachers cross-marking or check-marking groups other than their own.

Most of the science teachers considered that the main purposes of assessment were to provide information for inclusion on school reports, to rank students, to help students achieve success in the School Certificate, and to provide a basis for placing students in streamed or graded classes.

> One year I was Acting Head Teacher at the time that we were doing that, so I was in the thick of it, and it really made me see the value of doing common assessment - like well-structured, but common assessment tasks. If you're really going to do any kind of banding or grading, you've got to have some kind of basis to do it on (CMST2).

Marks from the formal assessment tasks were often kept in a centralised mark book in the faculty and this was considered to be the most important form of assessment (i.e., summative assessment). However, all the teachers acknowledged that they conducted less formal assessment procedures. This was generally not recognised as 'assessment' but rather as normal teaching practice. Interestingly, in most cases marks were allocated for class quizzes, bookwork, homework tasks, extension problems, reading proficiency, and practical work but recorded in a separate mark book. While teachers recognised that this information provided an overview of student performance, it was used only to provide 'comments' on school reports.

Assessment was dominated by particular task types that could be marked consistently. There was a preponderance of tests using multiple-choice items, supported by a few short-answer questions. However, there was a realisation, and a concern, that while tests were often the most convenient way to obtain information that could satisfy administrative requirements, they had shortcomings from an educational perspective. One teacher summarised the conflict as:

> I wanted to get away from this, 'Oh just give them a topic test.' Like I said, sometimes that's appropriate, and you want to do that; you want to have a bit of an idea about how kids can think, and what they know and that sort of stuff. Sometimes a test is an easy way to do it, but as you say, it's a little bit limited in what sort of outcomes it measures or tests (CVST1).

Other tasks often involved a large amount of time in marking, for example, the prescribed Student Research Project (SRP), which is compulsory in NSW for Stages 4 and 5. The workload was particularly onerous if one teacher completed marking across the cohort.

> At Year 10 level, one teacher makes all of Year 10 SRP, which is a huge job. We've made that decision basically, because the Board of Studies...you've got to be able to compare across the year with Year 10, so we divvy up the work - the person who marks the ISRP doesn't mark anything else for Year 10 (CMHT3).

Teachers at several schools expressed a desire to improve their assessment tasks while still achieving consistency and value out of the tasks. Teachers were keen to receive feedback about their tasks and seemed prepared to adapt these if necessary. A classroom teacher put forward the view that:

> We really need to have feedback on each assessment task, and make sure that we're sort of discussing how we can improve it, or whether it's worthwhile. Marking's always an issue, isn't it? We want things that are worthwhile for the kids to do, you can get a consistent marking scheme, and you can get a mark that's meaningful from it (CMST2).

The focus of assessment revealed in these baseline data was *assessment of learning* as opposed to *assessment for learning*, despite the NSW Board of Studies' current syllabus documents prescribing the opposite. Teachers were far more confident employing assessment in a summative way than in a formative way and lacked confidence in methods of assessment other than formal pen and paper tests. Overall, assessment appeared to be oriented towards the syllabus and there were no comments to suggest that there was a developmental influence in assessment in science.

Mid-project data

Science teachers demonstrated a clear understanding that the SOLO model was based on the cognitive development of students. They appeared confident when coding student responses and could communicate with each other using SOLO terms such as 'cycles', and the 'unistructural', 'multistructural' and 'relational' levels of understanding.

They displayed a willingness to apply information gained during the professional development sessions to both their assessment and teaching practices. Critically, teachers recognised that assessment strategies could be diversified and improved through the application of the SOLO model. Furthermore, they stated that once students' levels of functioning were ascertained, teaching could be modified to enhance learning.

> It's an assessment tool...where you are trying to find out…a couple of things…first it takes into account where a student's cognitive level, development's at. Trying to maybe in some degree, trying to be less quantitative and more qualitative, as to what a student knows. We see it as a tool to teaching as well as assessment, and as a framework for learning (CLST1).

When discussing assessment, all science teachers demonstrated an increased awareness of the value of formative assessment, and were confident in articulating the value of assessment as a learning tool. This was the first demonstration of *assessment for learning* being articulated with this science group of teachers. There was strong adoption of the need to ascertain the developmental situation for students, and to modify instruction and assessment to cater for the diversity of student learning needs. The use of pre-testing to assess prior knowledge had been implemented and was described by several teachers during these interviews:

> We use pre-testing and all that sort of stuff at the moment, which is very easily done under SOLO concepts (CMST1).

> I think the first thing I would be saying, using my SOLO eyes, is you need to know where these kids are to start with, before you can - you see, you'd probably need some sort of formative assessment, or even preliminary assessment if you want to call it that, just to find out from what basis you're coming off (CLST1).

There were references to assessment being used for different purposes, which had not been evident in the baseline data. For example:

> The nature of assessment can be different. That it can actually … well from my experience … it can actually unlock the key to knowledge for some, but

> not all kids. I think that in future, once we expose kids to it more – to different types of assessment tasks ... so the open-ended nature … that we'll get better results from it (CLHT3).

One teacher expressed the view that he would:

> Probably like to see it start to play more into using SOLO - less and less concentrating on it as an assessment tool. We're starting to see its value for learning in the classroom. That's what we're seeing it as. That makes sense, because assessment shouldn't be something that's different to learning (CLST1).

The experiences during the project had led some teachers to review and in some cases change their assessment practice. During the interviews many teachers described tasks they had set on multiple levels to remove any ceiling or limiting effects on student performance. Interestingly, teachers realized that they had been unaware of this as a factor prior to the activities experienced during the project. Almost all the teachers stated that they now included more open-ended tasks in their teaching and assessment:

> You're giving scope for people to answer in a more broader sense [*sic*], and to bring in other things, rather than limiting it, say like extreme example, multiple-choice tests, where you have no information about the thinking or anything that's going on behind the scenes, and you don't even know whether they guessed or not (CVHT3).

It was appreciated that SOLO provided a mechanism for addressing learning without marks, and allowed students to demonstrate their misconceptions. Teachers recognised the value of this as:

> I'm not putting someone down; I'm allowing people to progress, and it gives me a nice easy scheme of marking and addressing kids (CMST1).

> We're saying: 'Just tell me what you know', and it's probably a less-threatening way to them, and they are able to tell you what they do know (CLST1).

This perception represented a definite move away from assessment for summative purposes, which was so prevalent in the baseline data. In addition, attempts had been made to add a SOLO perspective to summative assessment tasks.

> We've tried to implement SOLO-type questions into assessment tasks. I've been Head Teacher here for 6 years, and our School Certificate assessment prior to that was just 150 multiple-choice questions. We've tried to adapt more of the short-answer and longer-answer questions into our assessment, and as part of that we've tried to put in some of these SOLO-type questions (CVHT3).

In terms of teaching, there was a view that learning about the SOLO model was worthwhile and relevant. It appeared to affirm much of what teachers did intuitively and had a revitalising impact on their teaching and their other roles within the school.

> I guess it's infiltrated the way in which I think about teaching – in terms of management of staff and personnel – it's also influenced my structure of testing items and questions with my own teaching, the fact that sometimes I can teach a concept, well not teach a concept, but try to get them to a relational understanding of the concept – and sometimes you won't get 100% of the class there (CDHT2).

And others:

> Not only is it just about assessment, but it's about the way you do things to do the assessment in the first place. So if you're thinking about just purely as an assessment tool, it's a lot more than that, because it's informing what you're actually doing, so that you can do that assessment. So it's sort of feeding back into your teaching, so that sort of connection would have to be made clear (CVST1).

> I can say: 'You're at this level, you're at that level, you're at R_1, U_1 – if you want to go to the next level, this is what you've got to do' (CMST1).

At a broader level, programming changes have been made to accommodate a SOLO focus in teaching at Crystal Lakes High School:

> We've taken a lot of this stuff, and put it into our programme, and we're starting to teach along these lines. Kids need to have some 'basic knowledge' first (CLST1).

Clearly, there was a major change in teacher focus about assessment between the baseline and mid-project interviews. The majority of science teachers articulated the importance of formative assessment as a means of monitoring student progress on a day-to-day basis, thereby articulating the *assessment for learning* agenda. They were enthusiastic to implement changes in their assessment tasks as long as they received constructive feedback from the research team.

Science Case Study: Charles (Head Teacher)

Charles had been teaching for eleven years. He was very enthusiastic and prepared to try different strategies within the faculty to improve student engagement in science. During the baseline interviews, he articulated that the focus of his 7-10 assessment was to match assessment to Quality Teaching (Professional Support & Curriculum Directorate, 2003) principles and link it to the syllabus outcomes. Tasks were common across a cohort and had to be achievable by all students while still extending the more capable students. He was concerned that all learners had equitable access to resources at the school, and that responses represented a student's own work. Therefore, he required that all tasks be completed in one school period. Within the school he felt there was a need to motivate students to maximise their achievement and for the faculty to try to prevent what he called 'task avoidance' by students. The faculty helped with task design and the development of a marking rubric that was given to students prior to task completion. Each teacher marked his/her own class, based on this rubric and informal discussion among colleagues sometimes resulted in the rubric being slightly altered during marking, if anomalies were evident.

At this time, Charles rated his assessment skills as 'high', especially in the approach to developing and marking rubrics. He was concerned about outcomes-based assessment and the amount of content included in the syllabus.

> Outcomes-based assessment is great for teachers producing tasks because a specific goal is evident. For students the system is poor; there is no longer a concept of pass/fail and so human nature drives many students to achieve at a lower (less effort) level. In addition, the BOS has crammed too much content and too many wordy outcomes into the syllabus.

During the mid-project interviews, Charles demonstrated confidence in discussing the SOLO model and ideas about student cognitive development.

> My understanding of SOLO is that it's baseline neurological work – how the brain works, and how working memory works – working memory has to be able to deal with basic facts first, then it has to be able to compress them through re-use, to make more working memory space available, so that then the facts can be linked, and get the relationship, and the big picture eventually down the track.

He had appreciated greatly the opportunity to learn about the SOLO model, and he perceived many applications of it to both his teaching and assessment. Since becoming involved in the project Charles had communicated with the rest of the faculty, resulting in amendments to the faculty programme and assessment policy. Whereas previously these documents had been organised along grid lines based on Bloom's Taxonomy, they were now built on the framework provided by the SOLO model with increasing complexity.

> We introduce them to the Multiple Intelligences grid, which is now set against SOLO, instead of Bloom's. Now, across the grid, where it used to say Bloom's 'remembering' through to 'evaluating' or 'synthesising or whatever, we changed that bottom level to 'Basic Facts'. Then 'multiple facts', and then 'some linked facts', and then 'relational facts', and then to 'the big picture'.

> We make it compulsory in our programme that every teacher does the basic facts. So really, the tests build on the basic facts. Your assessment builds on the basic facts. If you've got those basic facts, then you can go on and do those big picture things with them. The kids that can't do things with them, they've still got the basic facts there. So again, it's linking it back to that SOLO idea: 'Get those facts firmly embedded in their brain', not by rote-learning but by repeated exposure from multiple directions.

Assessment tasks that the faculty deemed 'good quality tasks' had been retained but a number of different SOLO-type items had been added. One particularly interesting idea for Year 8 classes was to provide an assessment task that did not include a maximum mark. The students were rewarded by the amount of relevant material they included in their responses and how they linked their ideas together. There had been some surprising results from students at both ends of the ability spectrum.

> To date we've done Year 7, and that unlocked an achieving under-achiever - a girl really really bright, who was achieving pretty good, near the top of the class, but it was obvious just from casual observation that she wasn't achieving her personal best. And we gave the whole year the SOLO task, and she absolutely shone above the rest.
>
> And at the other end of the spectrum, the kid with really low ability, that would read a normal test and shut down because of the reading component, and not even just put pen to paper – some of them have actually engaged, because they can now write what they know. They're not limited by a set of questions. There's only one thing that directs them, so I think there was one boy who got 13 – we put a mark value on it – he got 13. To me, he got 13 goals. It's not 13 out of 100, or something like that, and I guess my feedback to him was – he said: 'How did I go?' I said: 'You did absolutely fantastic! You wrote down 13 facts, which showed that you learnt something. Well done! Good job!'

While the majority of these changes had occurred for Years 7-10, Charles had also implemented SOLO components with his HSC group. One example was the inclusion of 'reversibility' questions to promote deeper thinking in the students. In these questions, students were given the answer and asked to explain why this might have occurred using their scientific knowledge of a particular biological concept. He also introduced his students to the hierarchical pattern of knowledge (demonstrated by SOLO model) using language that was appropriate to the students.

> We've gone from the language we used in the SOLO project to one the kids understand. We talk about 'basic facts', 'relational facts' and 'big ideas' – making the links and going to the big concepts.
>
> I don't think I go around – I don't look at each kid and think: 'There's an R1; there's an M2'. There's no way I do that. But I do consider it when I'm teaching.
>
> The other thing that I've got from SOLO is that the nature of assessment can be different. That it can actually, well from my experience, it can actually unlock the key to knowledge for some, but not all kids. I think the SOLO-type task leads kids down that path to do that, or opens the opportunity.

Charles and his project colleague had brought their faculty colleagues into the discussions about the SOLO model, and achieved faculty support for implementing SOLO into their teaching and assessment policies and practices. Furthermore, Charles addressed the school Executive and provided a staff seminar on Gifted and Talented Students (GATS) with a focus on the SOLO model. Finally, at the Principal's instigation, Charles introduced the Regional Director of Education to the principles of the SOLO model and demonstrated its application to groups of interested students in the school. The Regional Director was impressed by Charles' explanations and communicated this in writing to the Principal of the school.

DISCUSSION

It is clear when the baseline and mid-project results are compared that mathematics and science teachers demonstrated a number of changes to their assessment and teaching practices over the course of the two-years of professional development. Mathematics teachers focussed on ways of asking questions so as to encourage students to explain 'how' they calculated a particular answer. Although they retained tests as their main assessment task, they were interested in exploring different types of questions with many teachers including open-ended questions. Similarly, science teachers set out to improve the quality of their summative assessment items so that they allowed a greater diversity of students to demonstrate their scientific understandings. They were more critical of previous assessment tasks and more deliberate in their approach to developing their assessment processes. However, in some schools science faculties began to initiate longer-term, larger-scale changes around restructuring their science programmes.

The case studies provide two examples of longitudinal changes in teacher practices. Barry demonstrated a fairly conservative approach to assessment and teaching in mathematics at the beginning of the project with an interest in trying new ideas with his students emerging during his mid-project interview. This included an awareness of formative assessment to enhance student learning and his own teaching. Similarly, Charles (as Head Teacher) was prepared to try different strategies with his science students and was prepared to be a 'risk-taker'. However, he moved beyond his own classroom by initiating changes in the science faculty based on his experiences in the project. This was an important outcome and was the reason for the inclusion of Head Teachers in the project initially. While individual development in teacher practice is critical, sustainability and long-term impact in NSW requires policy adaptation at a faculty level.

A key feature for teachers in the project was the incorporation of the SOLO model as a theoretical framework. Based on working memory and student cognitive development it helped to explain to these teachers why particular strategies worked in the classroom and why other activities did not enhance student learning in mathematics and science. It was clear from their comments that the model provided a theoretical perspective that many found missing from many educational professional development activities. For example, the notion of 'modes of thinking' played a critical part in teachers' changed practice. In particular, they realised that the provision of concrete materials and practical tasks at specific points in the learning process helped in developing greater student understanding. Furthermore, many teachers referred to 'moving students through levels' or 'shifting a student from one level to the next' by referring to changes they had implemented in their teaching activities. For the majority of teachers this involved a shift to focussing on student conceptual understanding not just knowledge acquisition. In some schools, teachers actually introduced their students to the language and expectations of the SOLO model so that students were better able to monitor their own learning progress.

Overall, it was evident that teachers had moved from an *assessment of learning* to an *assessment for learning* perspective. In making this shift however, it was clear that they still perceived these as two ends of a continuum rather than the latter incorporating the first. Additionally, a number of issues emerged around assessment that required further exploration during the final phase of the project. These included:

- The infiltration of the stage 6 (HSC) assessment practices into years 7-10 assessment was still identifiable during the mid-project interviews. Teachers acknowledged that extensive use of scaffolding in teaching and learning could limit students' understanding but were concerned that the prescribed meaning of the verbs in the Stage 6 syllabus would mean that students whose responses did not conform to the set definition and requirements would be disadvantaged when they eventually sat for the HSC examinations. Therefore, teachers continued to instruct students in the syllabus-defined meaning of these verbs and this was introduced during Stages 4 and 5.

- There were concerns about the need to provide students with opportunities to respond to open-ended items. Issues of problems with student literacy and a generalised reluctance on the part of many students in years 7-10 to compose and write extended response answers were seen as impediments to teaching particularly in science.

- Consistency in teacher judgement was still seen as an issue where open-ended responses were used. Again, teachers considered that the best way to address this and protect the interests of students was to assign the marking of a particular task to one person.

- Despite the impact of the SOLO model on teachers, reporting practices in schools had not changed. In most of the schools there was a school requirement that grades or numbers be entered onto student reports, and some teachers saw this as incompatible with the SOLO model. Others attempted to develop some compromise on the issue but the systemic requirements and external pressures were of concern to these teachers.

CONCLUSIONS

Assessment is clearly a major issue for teachers in NSW, particularly given the high-stakes agenda in place due to the HSC. Greater needs for accountability and justification in the senior years have impacted teacher practice in years 7-10. The results provided in this paper indicate that it is possible to help teachers refocus their assessment and teaching practices to better meet the needs of their younger students but this takes considerable time. For example, the differences between baseline and mid-project interviews articulated here emerged after eighteen months. Yet, this level of commitment is beyond the scope of most professional development providers given the current financial constraints.

The strength of the project to date is that it has provided groups of teachers with the opportunity to "talk about teaching practice – what they do everyday, and being able to reflect on that, and talk about it away from the school environment" (CLST1). Furthermore, it has given groups of teachers the ownership to focus on their own assessment needs given the context of their own working environments.

ACKNOWLEDGEMENTS

The study reported in this paper was supported by an Australian Research Council Linkage Grant: LP 0455320. We thank the NSW Department of Education and Training our industry partner for their contributions to this project. Also, our thanks go to the teachers from the six DET schools involved who gave their views so honestly and openly.

REFERENCES

Bell, B. & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education, 85*, 536-553.

Biggs, J. (1996) Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.

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. (1993). Formative and summative assessment by teachers. *Studies in Science Education*, *21*, 49-91.

Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, *5*(1), 7-74.

Board of Studies (2003). *Science 7-10 Syllabus*. Sydney, NSW: Board of Studies.

Case, R. (1992). *The Mind's Staircase: Exploring the conceptual underpinnings of children's thought and knowledge.* New Jersey: Laurence Erlbaum Assoc.

Collis, K., Jones, B., Sprod, T., Watson, J., & Fraser, S. (1998). Mapping Development in Student's Understanding of Vision using a Cognitive Structural Model. *International Journal of Science Education, 20* (1), 44-66.

Fischer, K.W. & Knight, C.C. (1990). Cognitive development in real children: Levels and variations. In B. Presseisen (Ed.), *Learning and thinking styles: Classroom interaction* (pp. 43-67). Washington, DC: National Education Association.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine.

Goodrum, D., Hackling, M., & Rennie, L. (2001). The status and quality of teaching and learning of science in Australian schools: A research report. Canberra, ACT: Department of Education, Training and Youth Affairs.

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112.

Koetz, D. (1998). Large-scale portfolio assessments in the US: Evidence pertaining to the quality of measurement. *Assessment in Education: Principles, Policy, and Practice*, *5*, 309-334.

Linn, R. L. (1990). Essentials of student assessment: From accountability to instructional aid. *Teachers College Record*, *91*(3), 422-436.

Morgan, C. & Watson, A. (2002). The interpretative nature of teachers' assessment of students' mathematics: Issues for equity. *Journal of Research in Mathematics Education*, *33*(2), 78-107.

Panizzon, D. (2003). Using a cognitive structural model to provide new insights into students' understandings of diffusion. *International Journal of Science Education, 25*(12), 1427-1450.

Piaget, J. (1954). *The Construction of reality in the child.* New York: Basic Books.

Pegg, J. (1997). Assessing students' understanding at the primary and secondary level in the mathematical sciences. In J. Izard & M. Stephens (Eds.), *Reshaping Assessment Practice: Assessment in the Mathematical Sciences Under Challenge* (pp. 368-385). Melbourne: Australian Council of Educational Research, ISBN 0 8 6431 127.

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

Pegg, J. & Davey, G. (1998) A synthesis of two models: Interpreting Student Understanding in Geometry. In R. Lehrer & C. Chazan, (Eds.). *Designing Learning Environments for Developing Understanding of Geometry and Space* (pp. 109-135). New Jersey: Lawrence Erlbaum.

Pellegrino, J.W., Chudowsky, N. & Glaser, R. (Eds.) (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

Professional Support & Curriculum Directorate, (2003). *Quality teaching in NSW public schools. Discussion paper*. Sydney, NSW: Author.

Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.

Strauss, A. L. & Cobin, J. (1998). *Basics of qualitative research – grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Stephens, M., Clarke, D.J., & Pavlou, M. (1994). Policy to practice: High stakes assessments as a catalyst for classroom change. In G. Bell, B. Wright, N. Leeson & J. Geake (Eds.), *Challenges in Mathematics Education: Constraints on Construction. Proceedings of the 17th annual conference of the Mathematics Education Research Group of Australasia* (pp. 571-580). Southern Cross University, Lismore, NSW: MERGA.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758-765.

Watson, J., Collis, K., Callingham, R., & Moritz, J. (1995). A Model for Assessing Higher Order Thinking in Statistics. *Educational Research and Evaluation, 1*(3), 247-275.

Wilson, M. & Sloane, K. (2000). From principles to practice: an embedded assessment system. *Applied Measurement in Education, 13*(2), 181-208.