1-1-1998

How schools in Western Australia are responding to the pressures of accountability in Mathematics education

Lynne Messenger
Edith Cowan University

Edith Cowan University

Copyright Warning

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

How Schools in Western Australia are Responding to the Pressures of Accountability in Mathematics Education.

Lynne Messenger **St. No. 4790219**

Edith Cowan University

Master of Education

Supervisor: Alistair McIntosh

Date: 29th July, 1998.

A thesis submitted in partial fulfilment of the requirements for the award of Master of Education.

ABSTRACT

Western Australian primary schools have responded to the Education Department's call for school accountability in many areas of the schools' functioning. This study focuses on the endeavours that schools have made to demonstrate their accountability in the subject area of Mathematics.

The document "Improving and Reporting Schools' Performance" was made available to all state primary schools at the beginning of the 1996 school year as part of the Student Outcome Statement package. However, only a small number of primary schools have used the Student Outcome Statements in Mathematics in the National Professional Development Program trial conducted by the Commonwealth Department of Employment, Education and Training and supported by the Education Department of Western Australia during 1994 and 1995. A few other schools have used the Student Outcome Statements in Mathematics as part of their own school development program. Part of this study tries to identify whether the differences between how the pilot (NDPD trial) schools, those schools which are using the Student Outcome Statements but were not pilot schools, and schools which have not used the Student Outcome Statements at all, have responded to the call for accountability in the area of mathematics.

The study also investigated whether there were differences between country and city schools, and whether the size of the school or the presence of a specialist mathematics teacher affected the schools' responses to accountability.

A questionnaire was sent to one hundred randomly selected schools and to all eleven schools who participated in the EDWA pilot study, asking the principal to identify the approaches that have been used to justify accountability and the methods used to assess the student in their school in the area of mathematics. The parties responsible within the school for making the decisions about how the school responded were investigated and how the school was informing the wider school community of any changes were also investigated.

Three schools which seemed, from their responses to the questionnaire, to have approached accountability in innovative ways were approached and informal interviews were conducted to investigate further the methods of assessment and accountability and the processes that had been used within the school in the decision making process that had been undertaken.

Differences were found, in the ways that schools assessed their students and responded to accountability, between the pilot and the non-pilot schools and between those that were using the Student Outcome Statements and those which were not.

DECLARATION

I certify that this project does not, to the best of my knowledge and belief:

(i) incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;

(ii) contain any material previously published or written by another person except where due reference is made in the text; or

(iii) contain any defamatory material.

Signature

Date 29-7-98

ACKNOWLEDGMENTS

In preparing and writing this thesis I have greatly appreciated the help, advice and guidance received from my supervisor, Alistair McIntosh.
Thank you to the three school Principals who trialed my survey and gave me invaluable feedback with regard to its structure and the pertinence of the questions that it contained.
Thank you also to all of the school Principals, and Coordinators of Mathematics who so kindly responded to my survey. In this time of change and uncertainty I realise that time is a limited commodity and the time you gave me as you thought about and completed the survey is greatly appreciated.
Special thanks are extended to the Principal, the school Maths Co-ordinator and the school Development Co-ordinator who gave more of their valuable time to elaborate on their responses in the survey and who answered further questions regarding the processes that their schools had been through to establish their current school systems. Their hospitality was also much appreciated.
My three children, Elise, Ashley and Nathan, deserve special thanks for their support and patience while I hogged the computer.
Last, but definitely not least, thank you Mum and Dad for your encouragement and support, especially when you amused the children so that I could work without interruption.

TABLE OF CONTENTS

LIST OF TABLES

Introduction

The background to the study

Early in 1996 all government schools in Western Australia were provided with a document called *Improving and Reporting Schools' Performance* (Education Department of Western Australia (EDWA), 1996a). This document outlines three levels of accountability: the performance of students, the performance of schools, and the performance of system services. The aim of the document is to give direction to the schools in assessing the second level of accountability, the performance of schools, to set in place a system of 'quality assurance'.

> Quality assurance involves the establishment of processes to improve, monitor and report publicly on the organisation's performance against predetermined objectives. Quality assurance also assumes a commitment by all employees to monitor, and where necessary, to improving their personal performance. (EDWA, 1996a, p. 1)

School monitoring and improvement is considered necessary in seven areas of school performance: student learning, focus on improvement, curriculum delivery, learning environment, managing staff, and communicating with the school community. The staff within each school are expected to assess their own performance in each of these seven areas according to pointers given on a continuum, ranking themselves as being undeveloped, developing, functioning, achieving or of excellence. This self-ranking is

then to be validated by the district superintendent according to evidence that is provided by the principal. The district superintendent reports the evaluation of the school to the Executive Directors of schools.

Although the schools are asked to rank themselves according to the pointers, they are not specifically told how they are expected to do this. For example, in the area of 'student learning', the schools are directed to make judgements about student performance concerning two questions. The first question asks the school whether information has been collected on the students' performance in each subject area and how regularly this is done. "Regular, valid and reliable school-level information has been collected on student performance relevant to all of the school's performance indicators" (EDWA, 1996a, Chart) scores an 'excellence' level, whilst "information on student performance and/or performance indicators does not exist" (EDWA, 1996a, Chart) indicates a level of 'undeveloped'. The other levels are also given specific descriptors.

The second question asks how this information is being used to make judgements about student performances. "The information enables comparisons over time and between like groups. Reliable judgements are made about the level of performance being attained by students across all performance indicators" (EDWA, 1996a, Chart) is ranked as 'excellence', whilst "Judgements about student performance do not exist" (EDWA, 1996a, Chart) scores a ranking of 'undeveloped'. Other levels between these are given their own descriptors.

These judgements concern information that is collected about the students but there is no indication of the type of information that 'should' be collected. The type of information collected by the schools to show their accountability is left to the discretion of the staff at each individual school.

This document is part of the 'outcome package' and is expected to be used in conjunction with the various subject Student Outcome Statements by the schools as they make their judgements: "Schools will demonstrate educational accountability using the Student Outcome Statements" (EDWA, 1996b, p. 6). The specific subject information is contained in the individual subject 'outcome statements'. However, apart from a few schools which have piloted or otherwise elected to use the Student Outcome Statements, they will generally be unavailable to schools until 1998 when they have been reviewed and in some cases rearranged and parts rewritten. So for the present it is up to the individual schools to decide how their judgements on student performances are to be made. The desired movement, in how the schools respond to this challenge of being accountable, is towards the use of the various Outcome Statements (EDWA, 1996b).

The significance and purpose of the study

Although the primary schools in Western Australia have been given some guidance by the Education Department of Western Australia in the use of the *Improving and Reporting Schools' Performance* document and, in some cases, the use of the subject outcome statements, each school will inevitably respond in a different manner to these documents depending on the philosophies and experiences of the staff within the school, the documents that they have had the opportunity to come to terms with and in some cases the demands made by the individual superintendents.

It is important to look at the variables that will determine the school's response to accountability in one small area of the *Improving and Reporting Schools' Performance* (EDWA, 1996a), that is, the area of mathematics within the 'Student Learning' strand of the document.

Determining some of the variables that are affecting the schools' responses to accountability will give a better understanding of why schools have responded the way they have. It will also, perhaps, show some of the needs that the schools have in responding to the call for accountability in the way that is anticipated by the Student Outcome Statements.

It is hoped that, as a result of this study, the schools that participate will be given feedback on the different methods of responding to accountability in mathematics education which are discovered by the study, for their consideration as they continue to develop their own response.

Research questions

The main question that is being studied is to find how the state schools in Western Australia have responded to the pressures of accountability as outlined in *Improving and Reporting Schools' Performance* (EDWA, 1996a), in the area of mathematics education.

Part of this question it to determine whether the schools in the study have developed an approach to the assessment of mathematics that can be used to compare the students with their previous performances to monitor their development during their time at the school and the methods of assessment that will be used to monitor the students' development. Further to this, it is hoped to determine whether the approach made to accountability is the result of a combined effort by the staff of the school and whether there has been input from the parents and the extended school community.

There is an interest to find out how the schools which have worked with the Student Outcome Statements in mathematics have responded differently to those which have not, whether there is a difference between the schools situated in country

or city areas and whether the size of the school has had an affect on the responses made.

The research questions:

1. What are the differences in the methods chosen by schools that have and have not used the Student Outcome Statements to justify their accountability?
2. What are the differences in the ways that city and country schools have responded?
3. What are the differences in the responses of schools of different sizes (as determined by the number of teachers)?
4. Which groups within each school have had input into how accountability is demonstrated?
5. Which types of assessments are schools using to justify their accountability?
6. How are schools monitoring the progress of their students and which methods are they using to do this?
7. Are schools looking for alternative ways to demonstrate their accountability?

Review of Literature

Accountability

What accountability means

As part of the professional duty of a teacher, one is obligated to give feedback to three separate groups of people. The teacher's primary concern must be to inform their students of the educational aims of the programme, what is expected of them, and their individual progress in meeting these aims.

> That is, the student is an active participant in all aspects of the assessment process. Teacher-student interactions concerning assessment arise from a common understanding of the learning objectives and the work requirements, and the way in which each assessment strategy is to contribute to the developing picture of the students growth towards mathematical competence. (Clarke, 1988, p. 61)

Secondly, schools are obligated to report to the parents of the children on a regular basis. Parents need to know how their child is progressing, the skills that they are learning and the difficulties that they are experiencing.

Finally, teachers are required to report back to the educational system in which they work. The individual teacher would report to their own principal who in turn reports back to the educational authority who can then inform the government and in turn the community of what is being achieved. This latter form of reporting is that which is referred to as accountability. According to the Education Department of Western Australia:

> Accountability is the process through which schools demonstrate that they are performing effectively in terms of the education students receive. Schools must also demonstrate that they are operating within the policy framework of the government schools system. (1996a, p.15)

Ways that accountability can be determined

The accountability of a school system can be approached in two ways (Willis and Kissane, 1995a). The first method looks at the resources that are provided by the system for the students (inputs), in the form of materials, finances and personnel. The second looks at the products and achievements of the system (outputs).

In the past the output of the Western Australian school system has been done through the use of MSE (Monitoring Standards in Education) testing of a random sample of students in years 3, 7 and 10. An alternative to this are the Student Outcome Statements which are currently being developed in Western Australia and in many other education systems throughout the world. However, with the Student Outcome Statements, the emphasis is removed from what the children can and cannot do, as in the MSE testing, to viewing the development of every child as a continuum, and determining how far each child has progressed along this continuum in order to make decisions about the needs of the child to enable further progress. As Willis and Kissane state:

> The philosophy underpinning this approach to accountability is that desired student outcomes should be clearly articulated and it is these, rather than a plethora of policies and regulations about how schools should function, which should be a foundation for decisions about curriculum, teaching, assessment, professional development and so on. (1995a, p. 4)

However, according to Kissane (1997), MSE will continue to be used in Western Australia, although the items that it will contain are being based on the Student Outcome Statements rather than the curriculum objectives as has previous been the case. The items will therefore need to be responded to by children who have attained various levels of SOS. However, whether a student's response to an item will reflect their level of attainment on the SOS is very uncertain. He states: "A decision that a particular pupil has attained a particular outcome ought to be made on the basis of consistent observations, in a range of circumstances, over a period of time, and rely on the autonomous activity of the pupil" (Kissane, 1997, p. 80). Summaries of the findings of MSE will continue to be published as in the past.

Alongside this, each school is accountable to the regional superintendent who on an annual visit to the school requires the school staff to have assessed themselves against the pointers in the *Improving and Reporting Schools' Performance* (EDWA, 1996a) document. The school has to justify their self-assessment with their MIS (Management Information System), which is the data that the school has collected in order to monitor itself and make its judgements. This usually involves the collection of information related to student performance, this information is responded to with regard to further planning ie. identification of students at risk and future priority areas. Eventually these judgements, in the Student Learning area, will be based on the Student Outcome Statements, however, this form of accountability has been established whilst the SOS are only in use in a few schools, other schools will need to determine their own way of showing their accountability to the superintendent.

Ways that accountability is being approached

There is widespread use of 'outputs' in the form of 'outcome statements' or benchmarks as the basis for showing accountability. However, the way that accountability is determined varies considerably between different education systems.

The Development Program in Scotland (Madaus & Kellaghan, 1993, Howson, 1993) and the National Curriculum in England and Wales (Madaus & Kellaghan, 1993, Howson, 1993, Lofty, 1993) both use set 'attainment targets' for their students and conduct national testing at the key ages of 7, 11, 14 and 16 years. However, in Scotland these tests are used to validate the levels that the teachers have assessed their students at and are used as confirmation when reporting to parents, whereas in England and Wales the results are used to compare individuals and the results for all local schools are made public. As Madaus and Kellaghan comment about the system in England and Wales:

> The requirements that schools should publish aggregate assessment results to encourage free market competition within the system became at least as important as the more learning-oriented functions attributed to the assessments.
>
> In Scotland on the other hand, the use of the assessment results was restricted to providing diagnostic and formative information for the sole use of teachers, who could, if they chose, share it confidentially with parents. (1993. P. 462)

The results in New York State (Willis and Kissane, 1995b) are also for public scrutiny although accountability is assessed through a school review process where a team of 'experts' observe the educational practises within each school over a period of a week and then report their findings back to the school "to help to stimulate thinking about ongoing school development and improvement. The review process focuses on school climate, curriculum and assessment, teaching and learning experiences, and professional inquiry and development" (p. 15).

The Ontario Common Curriculum and Provincial Standards (Ministry of Education, 1993) is achieved by considering the individual in relation to the benchmarks (outcome statements). The main purpose of assessment is seen as being formative in nature, that is, to provide feedback to the student and to assist the teacher in making teaching decisions. A variety of assessment techniques are suggested. Ontario is split into local regions that have the autonomy to develop their own methods within this framework. Toronto is one such region.

The Toronto Benchmarks Program, (Larter & Donnelly, 1993) was established to give a common ground on which teachers can make their assessment of the progress of their students. The program relies solely on individual teacher judgements regarding the observed responses of students during the performance of open-ended tasks. "The philosophy behind the program is that instruction, learning and evaluation should occur simultaneously in the classroom on a continuous basis with many authentic performance activities" (Larter & Donnelly, 1993, p.59). There is no comparison of student attainments between schools but there is comparison of assessment techniques between teachers to provide moderation in order that a common standard can be set. "Unlike externally developed and scored tests,

Benchmarks allow teachers, students, and parents to collaborate and remain in control of learning and evaluation" (Larter & Donnelly, 1993, p. 62).

The New Standards Project in the USA (Simmons & Resnick, 1993), which is a joint project of 17 states and 6 leading school districts, bases its assessment on evidence of attainment contained in students' portfolios. The portfolio contains three kinds of students' work: "work chosen by the district, school, and/or students; prescribed projects and other extended activities; and responses generated by NSP matrix examination tasks" (p. 13). The portfolio is used as evidence of the students' level of achievement. Teachers are instrumental in the developments that are taking place. "We're talking about building an assessment system that heavily engages teachers in task development, scoring and using results to improve curriculum and instruction" (Simmons quoted in O'Neil, 1993, p. 18). The use of standardised testing has no place alongside the authentic assessments that are proposed. "The accountability system has to use those assessments that you're willing to have kids keep practicing ... To keep old tests while you're trying to bring the new system aboard would make teachers wonder which direction to go" (Resnick, quoted in O'Neil, 1993, p. 18). It is acknowledged that such changes to the ways that students are assessed requires a lot of teacher training and development.

Herman states that, "schools need support to implement new instructional strategies and to institute other changes to assure that all students can achieve the complex skills that these new assessments strive to represent" (1992, p. 77).

O'Neil (1992) asks, "But if an assessment is put into place to drive improved classroom instruction, should states implement such tests and begin reporting scores even if students have not had the opportunity to learn in ways measured by the tests?" (p.18). It could be argued that the accountability in this case is misdirected.

Outcome-Based Education.

In *Mathematics report: Report of the student outcome statements trial 1994-1995* (EDWA, 1996b), the conclusion with regard to accountability states:

> Student Outcome Statements will enhance teacher and school accountability through the collection and analysis of school information about achievement and through the implementation of a strategic plan to improve learning outcomes. (p.44)

The recommendations that are made continue:

1. Student Outcome Statements be adopted as a 'common standards framework' for school development planning across the Western Australian Government education system.
2. Student Outcome Statements be adopted as a 'common standards framework' for accountability across the Western Australian Government education system.
3. The link between school development planning, accountability and Student Outcome Statements be supported and articulated in all Education Department policy, guidelines and support material. (p.44)

The strong links that are being developed between accountability and the use of the Student Outcome Statements is evident.

What outcome-based education entails

A current trend amongst Western Education Departments and Ministries is the move to Outcome-Based Education. Outcome-Based Education generally requires a movement in the focus of education from "what are we doing?" and "how are we doing it, now?" to "what do we want to achieve, eventually?" and "what do we expect the children to be able to do, in the long term?" (Willis and Kissane, 1995b).

> An outcome is a culminating demonstration of the entire range of learning experiences and capabilities that underlie it. It occurs in a performance context that directly influences what it is and how it is carried out.
>
> The word 'based' means to direct, define, derive, determine, focus and organise what we do according to the substance and nature of the learning result that we want to have happen at the end.
>
> When we put these two words together, the term outcome-based implies that we will design and organise everything we do directly around the final intended learning demonstration. (Spady, 1993, p. 5-6)

Spady (1994) adds:

> [T]he demonstration must show significant learning; significant content is essential. Content alone, however, cannot be an outcome because it is inherently inert. Much like potential energy, it must be manifest through a demonstration process. (p. 18)

Three levels of student outcomes

General learner outcomes

The setting of outcomes starts with a long term plan of what the school expects all students to achieve by the end of their compulsory school years: these are known as 'exit outcomes' or 'general learner outcomes'. These long term outcomes do not focus on the subject content that is expected to be known by the students but on the competencies or life skills that the students are expected to be able to use. The focus of education is removed from the time that the student is expected to remain at school and placed on the standard of skill development that is expected before the student leaves school (Spady, 1993). These outcomes are general and not subject specific, focussing on the students' functioning within society and the abilities that the students should leave their schooling with. "They tend not to fit within the traditional subject areas in the curriculum and may address attitudes, appreciations and values, personal attributes, and work and process skills, in addition to 'academic' outcomes" (Willis & Kissane, 1995b, p. 2).

Redding (1992) states the five long term aims of the education system in Colorado. These are, that the students who graduate should be: "(1) self directed learners, (2) collaborative workers, (3) complex thinkers, (4) quality producers, and (5) community contributors" (p. 49).

General programme outcomes

General programme outcomes are more specific to the individual subject areas but still reflect the abilities that the students are expected to have before they complete their compulsory education. In many cases these "general programme outcomes [are] developed separately for each area of the curriculum, although they may share a common structure and purpose" (Willis and Kissane, 1995b, p. 3).

Specific programme outcomes

The most detailed outcomes are the specific programme outcomes. Rather than being the final outcomes of the entire program, these outcomes outline the expected changes that are to be identified in the students as they progress in each area of learning. These too are developed independently for each area of the curriculum. "The specific programme outcomes describe the outcomes that the system expects 'along the way' to the general programme outcomes" (Willis and Kissane, 1995b, p. 3).

The specific programme outcomes are arranged into levels of expectation so that the students can be seen to progress from one level to the next as they develop through their schooling.

In some school systems the specific programme outcomes are linked to specific grade levels. One implication of this is that individual students cannot progress from one grade to the next until they have attained the outcomes.

In other school systems the specific programme outcomes are arranged into a continuum that is not directly linked to grade levels to allow for individual differences. The implication of this is that there could be students working to attain several different levels within the same classroom. In fact, it is not unlikely that a student could be functioning at different levels even within the one subject area. This does not

mean that the teacher has to prepare different lessons for groups of students depending on their level of attainment, but rather that the activities should allow for students to be working at different levels with different expectations. A student who has shown to consistently work at a specific programme outcome level can be said to have attained that level, however it does not mean to imply that the student will always be able to operate at that level, this can depend on outside circumstances.

There are noticeable links between these programme outcomes and quality assurance as indicated in the *Improving and reporting schools' performance* (EDWA, 1996a) document, previously mentioned.

> Typically, education bodies describe their purpose in articulating programme outcomes as twofold. The first purpose is to develop a common language for talking about the results of student learning and hence to improve assessment and reporting. Thus the outcomes are regarded as part of the system's quality *assurance* mechanisms. The second purpose is to support improved teaching and learning by clarifying and making explicit agreed desirable outcomes of learning. In this case, the outcomes are regarded as part of the system's quality *enhancement* mechanisms. (Willis and Kissane, 1995b, p. 4)

Each independent School, Department or Ministry developing Outcome-Based Education seems to have developed its own balance between these levels of outcomes (Willis and Kissane, 1995b).

Outcome-based education and the curriculum

Outcomes do not constitute a curriculum in their own right, however, they serve to direct the curriculum. The outcomes can be seen as a framework which is used to guide the curriculum. "Curriculum content and structures should be planned to expand students' opportunities to achieve the outcomes" (Willis & Kissane, 1995a, p. 21). The outcome can be seen as the 'end' and the curriculum is the 'means to that end'.

Although the outcome statements do not dictate what is to be taught or how it is to be taught there are two considerations which those who expound the use of outcome-based education adhere to: these are clarity of focus and expanded learning opportunities.

Clarity of focus

When developing a curriculum to support the students' progression, the outcomes, both general and specific to the subject area, must remain at the centre of the professional decisions made by the teacher. It involves providing the students with the best possible opportunities to learn. These decisions are generally made at the school level and depend greatly on the expertise of the individual teacher.

Although the outcomes do not direct the teacher in the way that the students should be taught, some of the 'exit outcomes' such as qualities of self esteem, collaboration and cooperation imply that the learning situation should foster these attributes. As Willis and Kissane comment:

> Many educators also point to the importance of a supportive classroom environment for academic outcomes. This includes justice in access to resources, space and teacher time, and freedom from negative forms of anxiety brought on by excessive forms of competition and harassment such as teasing, sarcasm or remarks which stereotype or denigrate students or their efforts. (1995a, p. 24)

Not only is it important for the teacher to keep the bigger outcomes in mind but it is also seen as beneficial to the students to inform them of the outcomes so that they know what is expected of them.

Desirable outcomes can be seen to control the learning situation not only in the classroom environment, the types of activities provided, the encouragement given to the students to attain the outcomes, but also, in the way that these outcomes are assessed.

Expanded opportunities to learn

This consideration takes into account that students will inevitably learn at different rates and therefore it is assumed that most students need multiple opportunities to assimilate a given concept or skill.

Both Willis and Kissane (1995a) and Spady (1993) point out that this is not to be taken as meaning that mastery learning, where the curriculum is broken up into sections and the student needs to demonstrate a full understanding of one section before moving on to the next, and where the student repeats the same work until that

understanding is achieved, is being advocated. The intention is that a diverse range of learning and teaching styles be presented to the student so that learning is made possible and the expectation that if the student hasn't grasped the skill or concept on the first presentation of it then they have failed, is removed. In other words, the curriculum should not be governed by a time frame as in the semester, or the year. Nor should the curriculum be taught outcome by outcome but learning experiences should be developed so that the students have the opportunity to progress in a number of outcomes.

> OBE does not imply any particular pedagogical practices but the dual demand for clarity of focus and expanded opportunities to learn mean that teachers need a broad teaching repertoire...While a progression of learning should be evident in the learning experiences teachers provide, totally different activities are not necessarily required for students who have reached different stages in their progress. Often it will be in the questions teachers ask, the extensions they provide and what students do with the tasks that differences between students will be accommodated. Sometimes this will be planned in advance based on what teachers have already observed about their students. At other times, observing students as they engage in learning experiences will lead teachers and students to adjust things 'in action'. At still other times, teachers will decide to provide quite

> different learning experiences for students in the same class. (Willis and Kissane, 1995a, p. 28)

Three different philosophies of outcome-based education

Spady (1993) identifies three quite distinct types of outcome-based education. These he calls *Traditional, Transitional and Transformational.*

Traditional outcomes serve to make the existing curriculum more specific, resulting in the curriculum being split into small units of work. Spady (1993) suggests that these outcomes would be better called learning objectives as the "focus is primarily on skills and competencies" (p. 7). They are fairly narrow in scope, whereas true outcomes are usually broader in scope and not related to a specific learning activity and are not content dominated. "Because of their strong content grounding, these demonstrations are not generalizable across other areas of the curriculum or other performance contexts" (Spady, 1994, p. 19). 'Exit outcomes' are rarely considered and the outcomes do very little to facilitate the context of the learning environment.

Transitional outcomes are characterised by the use of 'exit outcomes' that look beyond the subject areas and yet there is still an adherence to the existing curriculum.

> It extends beyond the traditional form, as subject matter becomes more of a vehicle to assist in the cultivation and integration of higher order competencies. Critical thinking, problem solving, and effective communication skills, for example, are emphasised. (Spady, 1993, p. 8)

> Demonstrations are relatively complex and grounded in the kinds of competence that transcend given subject areas and that can be applied in a variety of relatively demanding performance contexts and settings. In this zone, demonstrations are generalizable across content areas and require substantial degrees of integration, synthesis, and functional application, thereby encouraging interdisciplinary approaches to developing the outcomes. (Spady, 1994, p. 19)

This form of outcome-based education does not seek to alter the curriculum as it exists but aims to improve the performance of teachers and the level that the students are working at.

> Most schools and districts have attempted to implement what Spady refers to as "transitional" OBE. Within this model, outcomes are identified for traditional content areas such as mathematics, science, history, and the language arts-as well as for less traditional areas such as community involvement, the ability to provide quality products, and the ability to work cooperatively and collaboratively with others. (Marzano, 1994, p. 44)

Transformational outcomes are aimed at challenging the whole structure of the way we see schooling. From the setting of the exit outcomes the knowledge and skills that

the students need to achieve are determined, and the curriculum is then designed to develop these. The conventional idea of subject areas is abandoned and the various subjects become tools to develop the 'exit outcomes'. Willis & Kissane (1995a) quote Spady:

> Mathematics would not be taught as a totally separate subject, but learnt in ways that link it to real-life problems, issues and challenges, so that it becomes the tool it was intended to be. Instead of teaching history by itself, we weave the evolution and historical development of ideas throughout everything we teach. In this way, students learn to thoroughly examine current issues and phenomena in depth and ask 'Why?' (p.9)

> Demonstrations require the highest degree of ownership, integration, synthesis, and functional application of prior learning because they must respond to the complexity of the real-life performance contexts. (Spady, 1994, p. 19)

In this highly integrated way the transformational curriculum is designed to develop the exit outcomes and create school leavers who are capable of dealing with life situations, both personal and in the work force, using current technology, being able to communicate effectively in various situations and capable of working independently or

collaborating as part of a team. "In short, the classroom becomes an active, high-challenge learning environment and performance center" (Spady, 1994, p. 22).

This form of outcome-based education has not been widely adopted. "Reformers have not heeded Spady's recommendations that to truly "transform" education, learning objectives within specific content domains must be discarded in lieu of objectives that reflect more realistic life roles" (Marzano, 1994, p. 44).

Outcome-Based Education in Western Australia

In 1987 the Education Department of Western Australia began a programme of devolution where the schools became responsible for their own curriculum and resources, however, it was seen that the schools needed "a common understanding of the curriculum requirements" (EDWA, 1996b, p. 1). The decision was made in 1990 to develop student outcome statements and this began in the areas of English and Mathematics.

> The outcome statements were intended to clarify for schools and teachers the elements of the curriculum for which they were to be accountable. They were also to provide clear guidance for teaching and a basis for reporting to parents on student progress (EDWA, 1996b, p. 1).

At the time other States and Territories in Australia were developing their own curriculum requirements and in 1993 it was decided that the States and Territories should collaborate on these under the auspices of the Australian Education Curriculum

Council (AEC). In the area of mathematics the document *A National Statement on Mathematics for Australian Schools* (AEC, 1993a) was produced, in which, major goals for school mathematics were set out. These goals reflect the long term aims of the system and so can be considered as 'general programme outcomes'.

A second document *Mathematics - a curriculum profile for Australian schools* (AEC, 1993b) was also produced, providing the curriculum framework that was suggested. This included outcome statements, pointers and work samples (to help teachers to identify when the outcomes had been achieved). Mathematics was divided into six strands each strand containing five to eight 'general learner outcomes'. 'specific learner outcomes' are divided into eight levels that are hoped to be achieved by year 10, although most students are expected to have attained level 6 by the end of their compulsory years of schooling. From this it can be seen that the Western Australian adoption of Student Outcome Statements reflects a transitional approach to Outcome-Based Education.

In the latter half of 1993 working groups from all interested parties in Western Australia undertook a review of the profiles, and with recommendations and changes made, these were published as the *Student outcome statements: 1994 working edition.* In the area of Mathematics very few changes were made. These outcome statements were trialed in some state schools during the years 1994-1995 to determine their usefulness and whether they are representative of the knowledge and skills that can be expected of students. Eighty eight schools trialed one or more of the eight curriculum area outcome statements. Thirty four schools specifically trialed the Student Outcome Statements in mathematics: this included primary and secondary and district high schools and "10 Kimberley schools who participated in the Inclusivity Trial" (EDWA, 1996b, p. 102). The inclusivity trial was conducted to consider the question: "Are the

Student Outcome Statements suitable for describing the achievement of students who may be at risk educationally while meeting the needs of all students?" (EDWA, 1996b, p. 67).

The results of the trials were published in a report (EDWA, 1996b) which views the outcome statements as a worthy addition to the curriculum framework and a vehicle for the betterment of education in Western Australia. "It has demonstrated their potential to improve learning outcomes for students" (EDWA, 1996b, p. 6).

After some minor changes to the Outcome Statements have been made, it is hoped that the Outcome Statements will be available in all Western Australian schools by 1998.

Assessment

It is claimed that the use of outcome statements as a guide to monitor the development of students does not require a change in the way that the curriculum is delivered to the students. "Often outcomes are promoted to schools with reassuring remarks that they do not necessitate changes to the curriculum albeit they necessitate changes to assessment and reporting" (Willis & Kissane, 1995b, p. 4). The terms in which the outcome statements and their pointers are expressed are more conducive to some methods of curriculum implementation than others. For example, these statements and pointers from the *Mathematics Student Outcome Statements with Pointers and Work Samples (working edition)* (EDWA, 1994) suggest that the student should be an active participant in the learning environment in collaboration with others: "comment upon their predictions in the light of results" (p.31), "inspect a polygon, prism or pyramid and tell a friend what sticks and joiners to collect to make the shape" (p. 26), "makes and tests conjectures" (p. 24), and "uses problem solving strategies that include those

based on selecting key information and representing it in models diagrams and lists" (p. 24). It follows that the type of assessment used should complement the teaching/learning situation. The *Mathematics Student Outcome Statements with Pointers and Work Samples* (EDWA, 1994), although not advocating any specific form of assessment, do provide the teacher with examples of students' work to show how the levels of attainment can be determined from classroom work.

Reasons for assessment

Assessment can generally be considered as the measure of a student's performance in a subject area. The assessment can be used: to help in planning the learning programme (these are known as formative evaluations), to measure an individual's achievement during the learning programme to monitor their progress, or at the end of the programme to measure the level of attainment (known as summative evaluations) (Collis, 1989). Assessment of a student's performance can also be used as a measure of the accountability of a system of education, a school or even an individual teacher.

The most commonly used forms of assessment

In the past, standardised tests and teacher constructed tests have been used extensively, and almost exclusively, to provide assessment information. Teacher intuition regarding the performance of a student has needed to be substantiated by the result of one of these pencil-and-paper type tests. Reporting, other than the teacher's comment, has usually been based on the scores achieved in such tests.

Standardised tests are those which have been commercially developed and trialed by a population of students said to be representative of the entire population of the region or country. The results of the trial are considered to be the "norm" for the

entire population and therefore a reference to compare other populations against. These tests have traditionally contained questions requiring short responses and in the United States have been predominantly multiple choice, where the correct answer is selected from a given range of possible solutions. Standardised tests have been considered of particular value as measures of accountability due to their norm-referenced nature. This reliance on standardised tests is currently being questioned. What are these tests actually assessing, and how useful are their results?

> Now, although the tests remain firmly established, influential decision makers, including well-informed politicians, are beginning to recognize their inadequacies. How useful for accountability purposes are data from norm-referenced multiple-choice tests when the tests compare students to one another rather that to established standards, and when they emphasize skills out of context rather than thoughtful application?" Educators who have long protested the misuse of standardized tests must concede that most of the tests that students take are devised by teachers, and that some of these are worse than the published ones. (Brandt, 1992, p. 2)

The majority of tests used in the classroom are teacher made. This, in some respects, can be seen as an advantage over standardised tests, in that the teacher has a good knowledge of what the students have been taught and should know. However, many

of these tests are very limited in the range and scope of question types and the type of responses that they elicit. Being a teacher, I can remember thinking, on many occasions, that certain students should have performed better on the test than they had. If a child can perform a certain task in the classroom as part of the learning programme, why are they not able to perform the task in the test situation? The answer that springs most readily to mind is that the test situation itself can create a feeling of anxiety in the student and this prevents the students from performing at an optimal level. Leder (1991) asks whether we should be surprised that mathematics tests often arouse strong emotions. The test itself can be seen as being divorced from reality. Very little mathematics is done in the real world under test conditions, so we should question whether it is right to assess students under test conditions. However, there is much pressure from the community and the students themselves to maintain these formal pencil-and-paper tests as the primary means of assessment in our schools (Leder, 1991).

The research of Clements and Ellerton (1995) highlights problems that children encounter in doing pencil-and-paper tests. They used an Extended Newman Error Analysis procedure to interview students, after the students had already taken the paper-and-pencil test. The students were asked to rework each question in turn and were then interviewed concerning their understanding of it. The students were asked to read the question aloud assessing their reading skills, retell the question in their own words assessing their comprehension of it. They were then asked to describe how they would go about solving the problem, assessing whether they could transform from the written format or diagram using an appropriate algorithm. Their processing skills were then assessed by whether the student could follow their plan through, and finally the student was asked to record their answer assessing their encoding skills.

In the study, Clements and Ellerton found that over one third of student responses on the test did not reflect the level of understanding found in the interview. Some students who had given a correct answer on the test were found to have either no understanding or partial understanding, whilst others who had given an incorrect answer on the test were found to have a full understanding or a partial understanding of the question. Clements and Ellerton used their results to "draw attention to the limitations of assessment via short answer and multiple-choice pencil-and-paper instruments" (1995, p. 12).

It can be argued that the use of these tests limits the curriculum to the type of items that are easy to test in this pencil-and paper mode, and that the style of the test inhibits the methods used in teaching. Nicholls et al. (1990) and Herman (1992) agree that testing will result in the quality of the teaching programme being negatively affected. "Accountability pressures encourage teachers and administrators to focus planning and instruction effort on test content and devote more and more time to preparing students to do well on the tests" (Herman, 1992, p.74). O'Neil concurs with this, and adds that many desirable learning outcomes are overlooked in the standardised test. This can be extended to implicate the formal teacher constructed tests too.

> [T]esting programs rarely succeeded in assessing students' thinking skills or their ability to synthesize content or solve problems. Moreover, educators felt pressured to "teach to the test" because high stakes were attached to test results....By slighting more complex student outcomes, critics say, basic skills

> testing essentially cemented into place low standards for student achievement. (O'Neil, 1992, p. 14-15)

Leder (1991) maintains that the formal test is a relic of the behaviourist paradigm as espoused by Thorndike where all learning was thought to be a bond created between a stimulus and a response. It was thought that by repetition this bond was strengthened and so the pedagogy of rote learning was established. As Leder comments: "In such a climate, the emphasis, of instruction and assessment is on the product" (1991, p. 118).

In opposition to the behaviourist paradigm is the view of the constructivist where knowledge is seen as being constructed by the learner. Each new experience results in the development of the student's understanding through the processes of assimilation (where the new learning is added to the existing understanding) and accommodation (where the understanding is modified to accept the new learning) (Chapman, 1972). For the constructivist the process of learning is more significant than the final product. If the processes of learning, that is how the student goes about constructing knowledge from the activity that is presented to them, is of primary importance, it can be seen that the formal pencil-and-paper test will do very little to assess the learning that has taken place.

The main advantages of the pencil-and-paper tests are that they are cost and time efficient in that they are easy to administer and mark. However, it is questionable whether tests are time efficient or whether they actually waste time in the amount of practice that is needed to prepare students for this kind of assessment.

It can be seen that pencil-and-paper assessments are incompatible with the Student Outcome Statements, where the focus is placed on the processes and 'expanded opportunities to learn'.

Assessments which are more conducive to outcome-based education

Wiggins, states that "typical tests...tend to over assess student "knowledge" and under assess student "know-how with knowledge"" (1992, p.27), he adds that tests should allow the student to display more of their intellectual processes in attaining a solution and suggests that the task given to students for assessment purposes should use the knowledge (or content) being tested as "a *necessary means* to a successful performance *end*" (1992, p.27). Tests need to be thought provoking and relevant (to a real world context or to student interests) and so be engaging for the student, developing the higher order thinking that is absent from standardised tests (Herman, 1992).

Higher order thinking is difficult to define, Resnick (1987) describes it as complex, involving multiple criteria which need to be structured. The judgements made are unique to each situation, hence a set procedure cannot be followed. There are often different routes that the thinking can take and often several possible solutions can result. There is often an element of uncertainty, and the thinker needs to be self-regulating, highlighting the desirable 'exit outcomes'. Higher order thinking requires effort.

Problem Solving and Investigations

In Mathematics, problem solving and investigations lend themselves well to the use of higher order thinking skills. "It is in problem solving that mathematical activity within the classroom most closely approaches the mathematical activity of the outside world. Because of this, problem solving activities provide unique assessment opportunities" (Clarke, 1988, p. 35). Schoenfeld (1985) mentions four aspects of understanding that can affect ability in problem solving, these are: the mathematical knowledge that the

student brings to the problem solving situation (Schoenfeld refers to this as the 'person's resources'), the known strategies that the student can employ to solve the problem ('heuristic'), the person's ability to access the relevant knowledge and strategies to make suitable decisions with regard to the problem (Schoenfeld calls this 'control' elsewhere this is referred to as metacognition), and the person's feeling towards solving problems and the subject of mathematics ('belief systems'). These areas of the cognitive structure give us a broader picture of the scope of assessment. However, this presents new problems for assessment as the thinking processes, decision making and beliefs are difficult to determine.

Szetela and Nicol (1992) suggest that special non-routine problem situations need to be created "that facilitate students' communication of their thinking" (p. 42). Many models have been developed for the assessment of problem solving (Szetela and Nicol, 1992, Garofalo and Lester, 1985, Clarke, 1988). A development of the problem solving situation used for assessment purposes can be seen in 'authentic assessments'.

Authentic and Performance Assessments

The use of problem solving situations can be seen in the call for more 'authentic' assessments. 'Authentic' assessments are those which complement the context of the classroom situation or the situation or context in which the outcome will be applied, rather than being divorced from the regular school experiences of the students.

There is a great need for the assessment tasks that are presented to students to show the students ability to apply skills and knowledge to real world situations and so demonstrate the students higher-order thinking skills. Marzano (1994) puts forward several reasons why these assessments should be presented to the students as a practical activity which is called a performance assessment, these are:

> (1) they provide clear guidelines for students about teacher expectations;
>
> (2) they reflect real-life challenges;
>
> (3) they make effective use of teacher judgement;
>
> (4) they allow for student differences in style and interests; and
>
> (5) they are more engaging than other forms of assessment. (p. 44)

Tasks selected to be used as performance tasks must have close links with the outcomes they are assessing. For example, if one of the outcomes concerns collaborative work then the assessment should be conducted with student collaboration and presented as a group effort. Likewise if a desirable outcome concerns the students as complex thinkers, the activity should be designed to include the students demonstration of their thinking processes, with detail to account for the processes that have been used and evidence to justify the inclusion of these.

Once a task has been developed, then a grading key for each of the outcomes being assessed needs to be developed, this is known as the rubrics for the assessment. Attainment of the outcome in its highest form would score, perhaps, a four, whilst little demonstration of the outcome would score a one. As performance tasks rarely have a single correct solution the rubrics take this into account by assessing the processes that have been used and the justification of these rather than the correctness of the solution.

> Tasks that have rubrics written specific to the proficiencies assessed can be scored reliably, whereas tasks whose rubrics are general cannot be scored reliably....Commonly, performance tasks are considered to have strong "face validity", which means that the assessment appears to measure what it is supposed to. (Marzano, 1994, p. 48)

Performance assessment tasks in England

The use of performance assessments was tried on a nationwide scale in England in 1991 (Nuttall, 1992). As part of the National Testing Program, which sets out to monitor the progress of all students in Grades 2, 6, 9 and 11, students in Grade 2 were given Standard Assessment Tasks that closely resembled the every day activities of the classroom and were administered and graded by the classroom teacher. Unlike most National testing, no standard instructions were given and the teachers were asked to introduce the activities, to individuals or small groups, in their own way. Although the tasks were demanding of teacher time and resources they were found to broaden the number of skills assessed, giving the teachers a better picture of the abilities of their students. The assessment tasks were enjoyed by most of the students who were not made "anxious" by the tasks. However, Nuttall added that "Performance Assessment can only take off with the will and expertise of the whole profession and with the trust of parents and politicians...despite all the care and effort, some will still not view it as rigorous enough" (1992, p. 57).

Madaus and Kellaghan (1993), however, describe the problems that were associated with the way that these assessment were presented to the teachers. These

included: the disruption of whole schools over a considerable period of time while some teachers were involved in individual and group assessments, the stress and pressure that the testing (which was used to compare teachers) produced in teachers whose classes were affected because they had to complete the testing and could not continue with their regular teaching program but had to hand their classes over to other teachers to be 'baby sat', and the huge cost in providing relief teachers to support the programme. There was also a lack of moderation in the way that the tasks were presented and assessed by teachers and, in general, teachers found that they didn't learn anything new about their students that had not been discovered in their own assessment programme.

The performance based assessments did not continue on a nation wide basis in later years but reverted back to pencil-and-paper assessments (Lofty, 1993) due to pressure from the country's Prime Minister. Although these authentic assessments had been intended to reflect the way that learning usually took place in the classroom, in reality this was not the case. Lofty (1993) lamented that the National Testing has been counterproductive to progress and improvement in the education system.

Assessments from Classroom Observation

The purpose of the National Testing Program in England was mainly for the purpose of system accountability. However, within the classroom, performance based assessments have a more viable place. Most teachers, from observations of what the children are doing whilst working in the classroom environment, adjust their teaching, the materials they introduce, the problems that they set and the direction of the learning. This is formative assessment at its most basic.

Teachers can formalise their classroom observations by setting out to assess their students as they work. This can be done by using a class list to check off desired behaviours or responses as they occur. Checklists can give a good picture of what the students were able to do in the classroom situation. They can be set up in two ways: firstly, with a list of the student names running down the page the desired behaviours can be placed across the top of the page so that a tick in the appropriate column can made quickly, alternatively, a separate page can be used for each child to record their strengths and attainments. Clarke (1988) suggests that when assessing problem solving in this way we must be careful to make sure that the assessments are multi-dimensional including the use of mathematical concepts, problem solving strategies, problem solving processes and also personal behaviours such as persistence, the level of participation, group skills etc. These could be awarded with a code eg. 1 - for little effort ..... 5 - for maximum effort, on the check list rather than just a tick. However, Clarke (cited in Herrington, Sparrow, Herrington & Oliver, 1997) warns that a checklist can be inhibiting if it becomes the teachers main focus.

An alternative observational record can be achieved through anecdotal comments. In this form of assessment, the teacher need only make a comment if an extraordinary, unexpected behaviour has been observed. "It really has to be something that is an unexpected insight or something that challenges a preconception they had about the child" (Clarke in Herrington et al., 1997, Anecdotal records - interview).

Whilst a teacher is observing the students as they work, well directed, open ended questions can help both the teacher and the student to clarify what the student is doing and their thinking process in doing it.

Portfolios

Wolf, LeMahieu and Eresh (1992) suggest the keeping of portfolios by students as a basis of their assessment, changing the focus of assessment from "measuring and reporting achievement for outside audiences... [to] ... encouraging students, teachers, and families to think hard about what is worth knowing and then making sure students know it" (Wolf et al., 1992, p.10).

The students keep folders of their work containing projects, journals and classwork. At the end of a period of learning they are encouraged (through discussion with the teacher) to select work, which is indicative of what they are capable of, to place in their portfolio. The student might be asked to justify the inclusion of each piece of work (Knight, 1992). This allows the child to reflect on the progress that has been made, and also allows the child to review what they have done (in reviewing earlier work a student might decide to add to it or change it due to new knowledge or skills).

The portfolio can be sent home for parental perusal and discussion or presented to parents by their children on a school portfolio evening (Hebert, 1992). They can be used in counselling the student (Wolf et al., 1992), especially in high school when subject or career decisions are being made. They can also be compared by teachers for assessment purposes, charting each student's growth as compared to the group (Wolf et al., 1992), teachers can act as a review board comparing the standards within their school (Wolf et al., 1992). If several schools use this method, several portfolios indicating different levels of achievement can be compared across the schools by the teachers, "[t]hey act as tough critics for one another, asking whether the curriculum is challenging enough, whether teachers are providing enough commentary to students, and whether substantial learning appears to be occurring

across the broad spectrum of students" (Wolf et al., 1992, p. 12). The portfolios can also be used by a committee, representative of the community, to assess accountability:

> They look at the quality of work, skill of teacher responses, appropriateness of grades , and whether all students have access to learning. The committee has three major purposes: (1) to publish a report of their findings and the implications of those data; (2) to design a long-term plan for how the district might move responsibly away from standardized testing; and (3) to think through how portfolio and performance task data can be responsibly used to inform planning, budgeting, and school change. (Wolf et al, 1992, p.13)

Advantages of using a portfolio are: students review, assess and reflect on the value of their own learning and as a result will be encouraged to strive to improve their efforts, to create variety in the student portfolios the teacher must vary the activities and assignments presented to the students (Knight, 1992), and students display qualities of "leadership and independence" that would go unnoticed with other forms of assessment (Hebert, 1992).

Reliability and Validity

Questions of reliability and validity arise with regard to all forms of assessment but Wiggins suggests that, "traditional views of validity and reliability need rethinking" (Brandt, 1992, p. 36).

A test's validity concerns whether the test result is an indicator of the knowledge, skills and abilities that are intended to be assessed. "To assess validity we compare our score on the test against an independent standard" (Collis, 1989, p.191). The main forms of validity used in the classroom are: face validity which is attained when someone who is "experienced" agrees that the test measures what it has set out to measure, "it is not in itself a totally reliable means of assessing validity" (Collis, 1989, p. 192), content validity which is obtained by comparing the test result to some other measure of the same skill, and predictive validity where a test is used to predict future performance (this is often the case when a test result is used to stream students), all of which are subjective measures. Factorial and construct validity are only used on standardised tests and rely on "sophisticated statistical techniques" (Collis, 1989, p. 192).

The test's reliability concerns whether a result obtained can be replicated. Conditions that affect a tests reliability include the wording of questions, consistency of scoring and time given, external conditions of the environment ie. light, noise, heat etc., the students well-being or their reading ability (Collis, 1989). Commercial tests are often trialed under varying conditions and statistical correlations performed on the results to assure reliability.

The validity and reliability of classroom based assessments are difficult to judge and although standardised tests are claimed to be reliable and valid, what it means is questionable when understanding is blatantly disregarded (Clements & Ellerton, 1995).

Herman queries that many people are developing alternative assessments but there is little data "on the technical quality of their assessments or about their integrity as measures of significant student learning" (1992, p.76).

Linn (cited in Herman, 1992) suggests that other criteria should also be used to judge whether assessments that are used to compare districts or schools are worthy. These include:

(a) determining the consequences of the assessment, that is, how those involved will respond to the assessment (pencil-and-paper tests have long been accused of creating anxiety in students);

(b) determining the fairness of the assessment, that is, does the assessment favour equally those from different backgrounds or sexes;

(c) the assessments transfer and generalisability (determining whether the assessment gives an accurate picture of the students' capabilities and whether it is correct to assume that these results are indicators of the range of abilities in other areas or schools),

(d) the cognitive complexity of the assessment should be determined, that is, does the assessment involve higher order thinking, Herman (1992) claims that from the appearance of an assessment it is difficult to tell how much high order thinking is needed as it has been shown that teaching to the test can reduce what could involve reasoning to be merely a rote solution;

(e) content quality, that is, any task given to students for assessment purposes should itself be worthy of solving;

(f) the content should cover the whole of the curriculum so that teachers are not tempted to omit the parts that are not to be tested;

(g) the meaningfulness of the assessment is important, they should be given in context resulting in "worthwhile educational experiences and in greater motivation for performance" (Herman, 1992, p. 76);

(h) and finally, the cost in terms of efficiency in time and effort on the parts of both the student and the scorer need to be considered.

The need for help to assist teachers in using alternative assessments

These alternative forms of assessment when used alongside Student Outcomes can give a better picture of the student's development and progress. The Outcome Statements become a "guide for learning not an endpoint to the learning process" (Diez and Moon, 1992, p. 38). "The essential idea underlying this innovation is that professional judgements should be based on the *outcomes*" (Willis, 1994, p.7). Willis states that a positive effect of using Outcomes is that the judgements of teachers are to be treated with respect. However, they will need to be able to justify these judgements for accountability.

It must be emphasised that where alternative assessments have been implemented, it has been found that the teachers need much training, guidance and support (Herman, 1992, Guskey, 1994). In fact, Vitali (cited in Guskey, 1994), in a study of teachers in Kentucky, after its new assessment program had been adopted, found that even though most teachers were very dedicated to their jobs and that they wanted the best for their students they "were very ill-prepared to adapt their instructional practices to the new demands of a more authentic, performance-based assessment program" (p.53). This was mainly attributed to the teachers' lack of background knowledge, training and experience concerning the performance based assessments. "Scattered, one-day staff development workshops"(p.53), are insufficient

preparation. The lack of appropriate teaching materials and the extra work involved were seen by teachers as problems, and so most teachers maintained their old teaching methods and were reluctant to change to the new assessment system.

Literature on methodology

The Questionnaire

The proposed study will take the form of a survey conducted by a questionnaire and case study. Englehart (1992) contends that the questionnaire lacks the rigours of alternative methods of research but goes on to acknowledge that "the survey has proved useful to social scientists in the study of social and social-psychological relationships including those relevant to formal education" (p. 292). Hillway (1969) agrees that the survey is useful "to determine present educational conditions and trends" (p. 31). Nisbet and Entwistle (1970) suggest that the questionnaire should be viewed as an "interview on paper" (p. 44).

The advantages and disadvantages of using a questionnaire

The main purpose of using the questionnaire is that the researcher can contact a greater number of respondents in a limited time and therefore can receive the answers and opinions of a greater number of subjects without additional effort. Nisbet and Entwistle (1970) state: "The obvious advantage in using questionnaires rather than interview is economy in cost, time and labour" (p. 44). The labour efficiency is evident not only in the collection of responses but also in the encoding of the responses during the data analysis phase of the study, as predetermined responses are often offered on the questionnaire in a multiple-choice format (Nisbet & Entwistle, 1970), whilst from

the interview the responses might not fall into specific categories, predetermined or otherwise.

One of the main disadvantages to the collection of information via a questionnaire is that it is impersonal (Hillway, 1996, Nisbet & Entwistle, 1970). Also, a good questionnaire is more difficult to construct than an interview, as the interviewer is available to clarify the meaning of questions whereas the questionnaire sender is not and so the questionnaire needs to be carefully worded so that there is no doubt of the intent of the question in the respondents mind (Hillway, 1996). Complex questions may also be answered superficially on the questionnaire, whereas in the interview situation further questions might be asked to elicit a more 'in depth' response. As the interview is prearranged, whereas the questionnaire is usually just posted, the recipient of the questionnaire has a greater liberty to dismiss it or reject it completely especially if they think themselves to be at a disadvantage (Nisbet and Entwistle, 1970). This can often lead to a number of recipients failing to respond. Nisbet and Entwistle (1970), consider it important for the researcher to try to find out why the recipients who do not respond chose to do so, and how these recipients are different to those who did respond. This, they believe, is frequently neglected.

Some of these problems associated with questionnaires can be overcome by the presentation and the wording of the questions that it contains. If the questions are carefully and unambiguously worded, the recipient is more likely to respond to it (Nisbet & Entwistle, 1970). The questions should also be unambiguously worded so that the respondent is clear about the intent of the question and so gives an acceptable response. Hillway (1969) suggests that the questionnaire should be "as short as possible but long enough to secure the information needed. A long questionnaire is often rejected simply because it takes too much time to read and answer" (p. 33). He

adds that it should appeal to the respondent and that questions should be specific and not evoke "superficial and unthinking answers" (p. 33) and that questions should not make the respondent uncomfortable as "some people would tell a lie rather than reveal a self-damaging truth" (p.33).

The types of questions

Multiple-choice questions, where a number of suggested responses are offered to be circled or ticked, are favoured by some researchers as they are easy to answer and simple to code and compare when the data is being analysed, however, the choices offered as the alternative responses must cover all possible responses as it is frustrating for the respondent not to find a response that applies to them (Nisbet & Entwistle, 1970). Sometimes the multiple-choice is in the form of a scaled item. Here the "individuals response is ... located on a scale of fixed alternatives" (Burns, 1990, p. 350). The alternatives given show the extent to which the respondent agrees or disagrees.

Questions where suggested answers are not specified but a space is given for the respondents to give their own answers are known as open-ended questions. Johnson (1977) suggests that "questions that are open-ended can be used to elicit ideas and opinions" (p. 152). For coding purposes the open-ended question must be fairly specific or the answers will be too difficult to classify. Nisbet and Entwistle (1970) suggest that "a multiple choice structure which includes the option 'Other (please specify)' may be a suitable compromise, though it is a sign of bad construction if this option is chosen at all frequently in the replies" (p. 46).

Sampling

In most studies it is difficult to reach the entire population that is being investigated due to the factors of cost and time, and so a representative group is often selected to be studied instead of attempting to access everyone in the population. "The investigator tries to select a sample that reflects the characteristics of the population as closely as possible, although the selected sample is rarely, if ever, a mirror image of the population" (Johnson, 1977, p. 139).

Sampling can best be conducted impartially by using a random sample. This involves the allocation of a number to every member in the population, deciding how many members of the population will be used in the sample and then using a table of random numbers, or calculator generated random numbers, to identify the members who will be selected for the sample. Each member of the population has an equal chance of being selected (Cohen and Manion, 1980). Cohen and Manion (1980) state that "one problem associated with this method is that a complete list of the population is needed and this is not always readily available" (p. 75).

The sample can be selected in a systematic rather than a random way eg. every fifth person listed. A stratified sample is obtained when the population is grouped according to special characteristics and a representative number of subjects is selected from each group. Quota sampling (Cohen and Manion, 1980) is similar to a stratified sample in that the researcher identifies groups within the sample but in this case a target is set for the number of representatives that are required in each group that is desired in the sample. Other forms of sampling that are discussed by Cohen and Manion are the convenience sample, where the nearest individuals are used to form the sample, purposive samples, where "the researcher hand picks the cases to be included" (p. 77), and snowball samples where "the researcher identifies a small number of

individuals who have the characteristics that he requires. These people are then used as informants to identify others who qualify for inclusion and these, in turn, identify yet others" (p. 77).

Johnson (1977) acknowledges that "a representative sample is by no means a general requirement in educational research. Students and teachers accessible to the investigator can be studied for the information they provide. The findings do not have to be automatically extrapolated to a larger group" (p. 140).

The size of the sample is a further consideration. Cohen and Manion (1980) suggest that the sample should not be less than 30 especially if statistical analysis of the results is to be performed successfully. They also suggest that the number of sub-groups and the comparisons that are to be made between these should be taken into consideration as well.

Sources of error

Johnson (1977) identifies three main sources of error when a questionnaire is used these are: sampling errors, non-response errors and systematic errors.

Sampling error is usually the result of chance and "not necessarily the result of mistakes made in the sampling procedures" (Cohen and Manion, 1980, p. 78). It occurs because the sample is only a representative of the entire population, some representative samples will more closely approximate the patterns in the population than others, there is no guarantee that the chosen sample is a good representative of the population or not. Hillway (1969) adds that sampling error "can be blamed on the selection of too small a sample" (p. 60).

Non-response errors can be the failure to complete a question. "Missing data are more frequent for items where the meaning is unclear, information is unavailable,

or information is of a sensitive nature" (Johnson, 1977, p. 163). A non-response error can also be the result of a failure to return the questionnaire, this can be due to the respondents disinterest in the subject, over work (not enough time to complete the questionnaire), or "they may believe the survey represents an invasion of privacy" (Johnson, 1977, p. 163). Hillway (1969) points out:

> The individuals to whom they are sent may actually constitute a truly representative sampling of the population; but the individuals who return them, unless virtually all respond, may actually be far from representative. (p. 60)

Johnson (1977) warns:

> While the number of nonrespondents may be known, the effects they have on the statistics of the survey can be difficult to judge. The magnitude of nonresponse error depends not only on the percentage of nonrespondents but the extent to which these individuals would have answered items differently to those who did respond. (p.163)

Nisbet and Entwistle (1970) feel that it is important to chase up the nonrespondents to find out why they did not respond and what their responses would have been. Hillway (1969) agrees that "those not responding or participating should be studied to

determine whether a bias of some sort exists" (p. 60). Johnson (1977) suggests that as the response rate becomes lower, the confidence that can be placed on the survey is also reduced. Englehart (1992) suggests that at least a 65 to 90 per cent return should be expected.

Systematic errors are those which produce bias in the survey and can result from the sampling, the questions asked (and the questions not asked) and through the respondents under- or over- stating their views.

The sampling can create the bias if a distinct group, with distinct ideals, within the population is left out, intentionally or otherwise.

The researcher must be careful to ensure that the questions asked are not 'leading' in some way or other. The respondent must remain unaware of the researcher's point of view. The way that a question is asked can provide a subtle bias that must be guarded against. Johnson (1977) suggests that to overcome this "the researcher can ask the respondents to consider the problem from several view points. The overall balance of the questionnaire can be neutral, both in the content of the questions and in the wording of the questions" (p. 152). He goes on to advise that a pilot study can be useful to overcome the problems inherent in the wording of the questions:

> Each person in the pilot study is asked to respond to all items on the questionnaire, perhaps noting any items where the meaning is unclear or where information is not easily available. The respondent might also be asked to suggest additional information that should be requested. (Johnson, 1977,p. 153)

The procedure for the collection of data

Hillway (1969) provided guidelines for the collection of data, these include that the questionnaire should be written in clear precise language that leaves no room for misunderstanding, the questions should be organised in a logical order, the better questions to ask are those that relate to fact rather than opinion and he suggests that the multiple choice questions should be primarily used. A covering letter should be sent with the questionnaire explaining the study and stressing its importance in a courteous and not demanding manner and giving clear instructions for the questionnaire's completion and advising of a date by which the return of the questionnaire would be appreciated. The inclusion of a stamped self-addressed envelope can help to encourage the return of the completed questionnaire.

The case study

Hillway (1969) describes a case study as:

> The intensive study of a single individual, several individuals, or a group at one particular point of time or over a period of time. It uncovers in detail what is true about an individual or group that may bear upon some phase of human behavior. Like those achieved in the typical survey, its results or conclusions are not so much prescriptive as descriptive. (P. 45)

The main advantage of the case study is that it "allows an investigation to retain the holistic and meaningful characteristics of real life events" (Burns, 1990, p.313).

The scope of what constitutes a case study is very wide and can include the use of various techniques such as observation, both structured and unstructured interviews and the analysis of documents (Burns, 1990).

One of the problems associated with the case study is the "subjective bias" that can occur. Burns (1990) states:

> The greatest concern has been the role of human subjectivity when selecting evidence to support or refute, or when choosing a particular explanation for the evidence found. It is easy for the case study investigator to allow equivocal evidence or personal views to influence the direction of the findings and the conclusions. (p. 325)

Another problem associated with the case study is that the findings are not easily generalisable to the entire population. Although, Burns does not consider this to be a problem: "Case studies are focused on circumstantial uniqueness and not on the obscurities of mass representation" (1990, p. 326).

Questions of reliability and validity frequently occur with case studies. Burns suggests that "to improve reliability and enable others to replicate your work, the steps and procedures must be clearly explicit and well documented in the final report (1990, p. 328). Validity is seen as a problem because of the lack of sampling techniques and standardised instruments of measurement. Burns (1990) suggests that to improve the construct validity multiple sources of evidence, and a chain of evidence that links parts together, should be used.

The analysis of questionnaire data, according to Burns (1990), "may simply consist of determining the frequencies for the major variables involved in the study" (p. 365). Where the exploration of relationships between variables is being studied crosstabulation of the variables and their frequencies is explored (Burns, 1990).

Case study data are commonly in the form of notes "and may have been derived from interviews, observations and documents.... These notes should be organised as an ongoing process so that as the study progresses the investigator has some sense of the direction it is going and the confirmations and contradictions that are arising" (Burns, 1990, p. 323). This, Burns suggests, is best done if the events in the data are coded to highlight patterns and similarities.

Theoretical Framework

A socioconstructivist view point is held to in this research. It is believed that children best learn when they manipulate their environment in collaboration with fellow students. During these processes the child is able to question their understanding and assimilate and accommodate new learning into their understanding so constructing their own knowledge. "Instructional activities should be designed to give rise to genuine mathematical problems for the students. Such problems constitute opportunities for them to reflect and reorganise their current ways of thinking" (Yackel, Cobb and Wood, 1992, p. 64). Interactive communication between students as they work is essential for as the students put their thoughts into words and test their ideas by sharing them with others their own understanding is clarified. The other students are at liberty, with reflection, to agree or disagree with the points of view that are being expressed. They can ask questions to clarify meanings and voice their own

opinions, in the hope that conflicting points of view are resolved (Yackel, Cobb and Wood, 1992). Group work where students can share experiences and discuss them are of paramount importance. "Constructing a reality of "shared" experiences by all concerned encourages the interactive communication that serves in encouraging the reorganization of experience as well as the autonomy of its participants" (Steffe, 1990, p. 394).

It is believed that all forms of assessment should be conducted as part of the regular learning experiences of the children, "every teacher is continuously offered a wealth of assessment information during the instructional process. Many act on this information, but few document it" (Clarke, Clarke and Lovitt, 1990, p. 120). Much can be learned from talking and listening to the children as they work and "classroom questioning offers possibly the best chance to monitor the development of meaningful understanding" (Clarke et al., 1990, p. 124). Every endeavour has been made to prevent these beliefs from influencing the questions presented in this research.

Method

The sample

There were two samples in the first part of the study. As there were only eleven schools that had piloted the Student Outcome Statements in mathematics, questionnaires were sent to each of the principals of these schools. The second sample consisted of 100 schools randomly selected from the list of primary and district high schools in *Western Australian Schools - Alphabetical List (Semester 1, 1997)* (EDWA, 1997a). The pilot schools already in the first sample were excluded from this sampling. The schools were each given a number according to their listing and a table

of random numbers was used to select the schools. The data from each sample was kept separate for the purpose of analysis. The data was analysed statistically and the two sets of data were used for a descriptive comparison.

The sample in the second part of the study consisted of three schools selected from the original samples who were contacted by telephone and agreed to discuss further their response to accountability. The principals of these schools had also indicated that they were using a method of assessing and/or recording student development that showed some form of innovation on the part of the school.

The design

The variables that were investigated consisted of whether the school had or had not used the *Mathematics Student Outcome Statements* (1994), whether the school was situated in the country or in the city, along with the size of the school to compare if there were differences in the way that each category had approached the monitoring of student development within the school.

Instruments

The initial instrument used in this study consisted of a questionnaire (see appendix 1). The questionnaire firstly determined the location and size of the school. Then the questions ascertained the general interest of the school, its areas of school development and whether there had been any interest in the area of mathematics especially in the use of the student outcome statements. The next section was designed to find out how the school had approached accountability in the area of mathematics, whether a unified school plan had been developed and if so who, within the school

community, had taken part in the development of this plan and how the students were assessed in relation to both the classroom and the school monitoring.

The second part of the study was in the form of unstructured interviews and the observation of school documentation in three of the schools who had developed an approach to student monitoring and were willing to participate when contacted by telephone.

Procedure

The questionnaire was pilot tested by three principals (who were not selected in the random or the pilot sample). They were asked to give feedback on the clarity of the questions, whether or not they could determine any bias in any of the questions, whether any of the questions could be interpreted as being inappropriate, the length of questionnaire and how they felt that the survey could be improved. They were also asked to give feedback as to whether they as a principal would be inclined to respond to the questionnaire if they had received it through the mail.

The questionnaire was posted to each of the schools with a stamped self-addressed envelope and an introductory letter (see appendix 2) setting out the purpose of the study and expressing appreciation of the completion and the return of the questionnaire. The schools which did not responded to the questionnaire by the due date were sent a letter of reminder and asked to either return the completed questionnaire or respond by indicating that they required another (see appendix 3).

For the second part of the study three of the schools who indicated that they were developing systems to monitor student progress in mathematics on the questionnaire were contacted, by telephone, and a suitable time for a visit to the school was arranged. The visit was in the form of an informal interview with the principal

and/or the teacher involved in developing the school plan, inorder to gain their comment on the processes that they had used in developing the plan and the advantages/ disadvantages they had found in using their plan. At two schools, the plan that had been developed was viewed, and discussion took place on how it was working to provide information about student development and to supply evidence to justify the level that they have given themselves in regard to *Improving and Reporting Schools' Performance* (EDWA, 1996a). Their observations and feelings regarding the implementation of accountability measures and student outcome statements in general was also sought. Written consent from the principal or teacher at these schools was obtained prior to any observations being made.

Data analysis

The computer program SPSS was used to help organise the data from the questionnaire.

The answers to questions 1 were to be used to classify each of the randomly selected schools according to size, however, the student numbers in each school from the *Western Australian Schools - Alphabetical List* (EDWA, 1997a) was more useful for this. The school location (city or country) was determined according to whether the school appeared in the schools' list in the *Metropolitan Telephone Directory* or not. This information provided the first two school categories.

The data from questions 2 and 3 was used to classify the schools as 'Whole School Using Outcome Statements in Mathematics', 'Some Teachers using the Outcome Statements in Mathematics' and 'Outcome Statements in Mathematics are Not in Use'. This gave a third school category.

The data from question 4 was used to compare the schools according to the school categories already outlined. These comparisons were presented as a frequency table.

Question 5 was included in the questionnaire to verify my assumption that all schools were aware of the *Improving and Reporting Schools' Performance* (EDWA, 1996a) document and that all schools were using it to determine their levels of accountability. A 100% 'Yes' response was expected to this question. Had any school respond 'No', the school's data would have needed to have been deleted from the analysis of question 10.

The data from question 6 was used to determine whether there had been any influence, due to the existence of a specialist mathematics teacher, on the adoption of the student outcome statements as indicated by the results of questions 2 and 3. A cross tabulation of the data occurred.

In each of questions 7, 8, 9, 10, 11 and 14 there was a likelihood that there would be multiple responses. The analysis of this data consisted of a frequency table. These frequencies were further broken down according to the three school categories.

The responses to questions 12 and 13 were seen as if on a continuum. Strongly agree was given a value of 5, Agree - 4, Undecided - 3, Disagree - 2 and Strongly Disagree - 1. The responses were tabulated and the means and standard deviations calculated.

The responses to question 15 were used to send feedback to the principals who had asked for it.

The 'further comments' helped in the selection of the three schools for the second part of the study. They also were discussed in the report as they helped to clarify why some of the schools responded as they did.

Similar analysis was conducted for the pilot study group and a verbal comparison between the two groups resulted.

A description of the school's procedures resulted from the information received in the second part of the study.

Limitations

The process of collecting the data for this study relied heavily on the cooperation of the schools, especially the principals. Nisbet and Entwistle commented: "In practice it may be very difficult to obtain a 70 per cent response rate from certain groups of people. Head teachers, college principals or managers in industry may well be too busy" (1970, p. 52-53). As the cooperation of the principals contacted in this study was crucial and the lack of it would have undermine the study, an introductory letter was sent with the questionnaire (see appendix 2), as the questionnaire was a fairly impersonal document. The number of nonrespondents was a concern especially in that it could cause a bias in the data if all nonrespondents would have answered the same way.

Of the 100 non-pilot schools, 56 responded to the survey and 8 of the 11 pilot schools (72.7%) also responded. Of the 44 schools that did not respond to the survey it was suspected that the majority of the schools would not have implemented the SOS in mathematics and therefore the data could be biased due to this. My suspicion is based on the fact that on receiving the survey and glancing at it, one principal rang to ask why he had been sent the survey as they were not using any SOS and therefore the survey did not apply to his school. If other principals had reacted in the same way for the same reasons it could be assumed that the data could be biased in favour of the use SOS.

Some of the questions may not have contained an appropriate response, even though it was endeavoured to cover all possibilities, and so the "Other (please specify)" response was included. If there had been a great use of this category there would have been great cause for concern, however, the only question where this category was greatly used was in question 14 where the alternative response that was frequently given was 'School Based Decision Making Group', this would have been an appropriate category to have included in this question.

Two other problems that were found with the questionnaire are, firstly, in questions 9 and 10 a box for 'none of the above' would have been helpful as many schools did not respond to the question and the school making a comment had to be relied on, otherwise no response had to be assumed. Also, in question 11 some schools ticked just the 'all teaching staff' whilst others ticked 'some teaching staff' as well as the 'all teaching staff' response. Whilst encoding the data it was necessary to credit an affirmative response to 'some teaching staff' only where an 'all teaching staff' response was not made as well - giving a category of 'some teaching staff only'.

Nisbet and Entwistle (1970) warned that there are serious limitations in trying to infer causality and in extrapolating the data from the sample to the entire population these were guarded against in drawing conclusions from the research.

As many schools had not embarked on the use of SOS in mathematics, the selection of three schools for the second part of the study was limited.

Results

Feedback from the pilot study.

The comments received from the three participating principals was positive. All three principals considered that the questions were appropriate and that the intention in each question was clear. The length of the survey was also considered to be "fine" and "not a problem" although the third principal thought that if he had stopped to write what he had wanted to comment rather than just ticking boxes he would have taken considerably longer.

One principal did comment that there was a possibility of researcher bias being evident in question 6 where it was asked whether there was a specialist mathematics teacher in the school, in question 7 which asked for the comment that best described the school's approach to accountability and in question 12 which asked if the *Improving and Reporting School's Performance* document had increased the communication between members of the teaching staff. The principal, however, could not elaborate on this feeling.

Another principal informed that there had been a sequel to *Improving and Reporting School's Performance* (EDWA, 1996a) called *School Performance: A Framework for Improving and Reporting* (EDWA, 1997b). This information was used to restructure question 5 of the survey and the letter to the principals.

Data from survey.

Size of schools

The size of the school was difficult to determine from the information given in Question 1 of the survey (showing the numbers of full and part time teachers on the staff) especially as some schools hadn't answered the question. It was decided to use the number of students in each school to indicate the size. The numbers of students were obtained from the Education Department publication *Western Australian Schools - Alphabetical List* (1997a). This information was grouped in lots of 100 students, as shown in table 1: The sizes of the non-pilot schools as determined by student numbers; and table 2: The sizes of the pilot schools as determined by student numbers.

Table 1
The sizes of the non-pilot schools as determined by student numbers.

School size	Frequency	Percentage
0 - 99	9	16.1
100 - 199	8	14.3
200 - 299	11	19.6
300 - 399	11	19.6
400 - 499	10	17.9
500 - 599	4	7.1
600 - 699	0	0.0
700 - 799	1	1.8
800 - 899	2	3.6

Table 2
The sizes of the pilot schools as determined by student numbers.

School size	Frequency	Percentage
0 - 99	1	12.5
100 - 199	2	25.0
200 - 299	1	12.5
300 - 399	1	12.5
400 - 499	1	12.5
500 - 599	1	12.5
600 - 699	0	0
700 - 799	0	0
800 - 899	0	0
900 - 999	0	0
1000 - 1099	0	0
1100 - 1199	0	0
1200 - 1299	1	12.5

Location of schools

Whether or not the school was classified as a city school or a country school was dependent on whether the school was listed in the Metropolitan Telephone Directory (White Pages - Perth, Telstra, 1997). Of the non-pilot schools, which responded, 26 (46.4%) were located in the city and 30 (53.6%) were located in the country (ie. outside the metropolitan area). For the pilot schools there were 4 (50%) in the city and 4 (50%) in the country.

Use of Student Outcome Statements

Of the 56 non-pilot schools, exactly 50% of the schools had elected to implement at least one of the eight curriculum area Outcome Statements. Some of these schools were using several curriculum areas of SOS as shown in table 3: Number of SOS learning areas being used in the non- pilot schools.

Table 3
Number of SOS learning areas being used in the non- pilot schools.

Number of SOS areas	Number of schools	Percentage of schools
0	28	50
1	8	14.3
2	10	17.9
3	2	3.6
4	1	1.8
5	1	1.8
6	3	5.4
7	2	3.6
8	1	1.8

Of the schools that have adopted none of the curriculum areas of the SOS one school commented that the Outcome Statements were "not available to any but pilot schools", four other schools indicated that they were waiting for the final documents to be published before proceeding and one school was "waiting for some guidance from EDWA". Another school stated that in previous years there had been work on the SOSs but due to the State School Teachers' Union ban for 1997, this work had "been suspended". At two other schools their comments indicated that substantial planning had taken place although implementation had not as yet occurred.

One of the schools that had implemented several SOS learning areas stated:

> "We have many graduates who understood the concept of SOS - In a meeting at the commencement of the year we determined our data collection methods and the teachers agreed to use the SOS in the above areas (Arts, English, Mathematics, Science, Health and Physical Education, and Technology and Enterprise). Teachers

> are comfortable in the use of outcome statements. They use them for programming and planning and assessment."

At another school, where Outcomes were being used in several learning areas, the following comment was made: "SOS are used extensively in our portfolios when reporting to parents. For whole school MIS, SOS were used in Maths, Science, The Arts and Health and Phys Ed."

At one school where the staff were using just one SOS learning area, the comment explained the difficulty that they had had in obtaining help in this area from EDWA.

Mathematics was found to be the most used curriculum area of SOS being adopted by whole schools, with Technology and Enterprise being the second most used. The results are shown in table 4: Number of non-pilot schools electing to use each of the eight SOS curriculum areas throughout their school.

Table 4
Number of non-pilot schools electing to use each of the eight SOS curriculum areas throughout their school.

Curriculum area	Whole School
The Arts	7
English	13
Mathematics	17
Science	11
Health & Physical Education	11
Society & Environment	5
Languages other than English	5
Technology and Enterprise	14

Some additional interest is being shown by individual teachers where there has not been whole school adoption of the curriculum area at present.

It was commented that one teacher on transferring from a pilot school had continued to use the SOS in that area.

Table 5
Number of SOS learning areas being used in the mathematics pilot schools.

Number of areas	Number of schools	Percentage of schools
0	1	12.5
1	1	12.5
2	1	12.5
3	1	12.5
4	1	12.5
5	0	0
6	0	0
7	2	25.0
8	1	12.5

The number of SOS learning areas that have been adopted throughout each of the pilot schools is shown in table 5: Number of SOS learning areas being used in the mathematics pilot schools. The learning areas that have been adopted are shown in table 6: Number of pilot schools electing to use each of the eight SOS curriculum areas throughout their school.

It is interesting to note that one of the pilot schools in the Mathematics area is not using any of the SOSs. According to a contact in EDWA this was due to the large staff turnover that had occurred within the school.

Table 6
Number of pilot schools electing to use each of the eight SOS curriculum areas throughout their school.

Curriculum area	Whole School
The Arts	3
English	4
Mathematics	7
Science	4
Health & Physical Education	4
Society & Environment	4
Languages other than English	1
Technology and Enterprise	5

It is not surprising that as these were the pilot schools in mathematics, mathematics was the most used curriculum area. As with the non-pilot schools Technology and Enterprise was the next most widely adopted learning area.

Along with information obtained from question 3 of the survey showing additional interest by some teachers in each of the curriculum areas, a third category of school was determined these are shown in table 7: Use of SOS in mathematics in non-pilot schools, and table 8: Use of SOS in mathematics in pilot schools.

Table 7
Use of SOS in mathematics in non-pilot schools.

Category of use of SOS in maths.	No. of school	Percentage
Using SOS throughout	17	30.4
Use of SOS by some teachers	4	7.1
SOS not in use	35	62.5

Table 8
Use of SOS in mathematics in pilot schools.

Category of use of SOS in maths.	No. of school	Percentage
Using SOS throughout	7	87.5
Use of SOS by some teachers	0	0
SOS not in use	1	12.5

Proposed introduction of Mathematics SOS

When each of the non-pilot schools are aiming to introduce the mathematics SOS into their school, if they have not already done so, is shown in table 9: Proposed year of adoption of the SOS in mathematics in non-pilot schools.

Table 9
Proposed year of adoption of the SOS in Mathematics in non-pilot schools.

Year	Number	Percentage
Already in use	11	19.6
1998	9	16.1
1999	3	5.4
Has not been decided	33	58.9

It is interesting to note that of the 17 schools that responded positively to the Student Outcome Statements in the Learning area of Mathematics being developed throughout the school only eleven of these stated that the Mathematics Outcomes were already in use. Of the other six schools two said that they intend to implement them in 1998 and four stated that it had not been decided. This discrepancy could lead to the assumption that there is a problem with the reliability of the data, although, as one school commented: "We have been using a draft version at this stage. Hope to revise in 1998" could imply that there was some confusion as to whether I was asking about the trial document or the official one that is yet to be distributed, especially for the two schools that have said that they are using SOS in Mathematics but intend to implement it in 1998. Alternatively, from other comments such as: "Number in 1998 Spatial Knowledge and Measurement in 1999", "Starting small! ie part of SOS", "At this stage Yr 6 & 7 will adopt the use of SOS in Maths at this school in 1998 - rest not decided", "Start slowly - carefully and plan effectively", "We have focussed mainly on

the number and Working Mathematically strands this year. Next year we will extend to another area determined by the MIS", and from further discussions, I have found that most of the implementation is in a partial form at this present time (including pilot schools) and so some schools could be assuming full implementation as being the object of my question, thus resulting in the discrepancy .

Four schools commented that 1998 would be a planning year and so no decisions had yet been made. Three other schools, that had made no decision as to the time of implementation, displayed negative feelings towards SOS: "When there is a comprehensive inservicing of these outcome statements. When outcome statements are fully resourced", "Massive workload will begin when necessary" and "Staff are reluctant to implement anything until they have to. This is the result of jumping in early on previous occasions, only to have the rules changed after much time and effort has been expended". One school had attempted to implement the Math SOS but there had been lack of support so the decision had been deferred. Another school where the decision of timing had not been made commented more favourably: "Some are tasting the maths at the moment".

A school where the Math SOS are to be implemented in 1998 said, "Very strong commitment to outcomes based learning. A strand will definitely be a focus next year. We will focus on Working Mathematically as a starting point".

A school where the Math SOS are in use commented that "Open ended tasks and moderation of these helped develop a thorough working knowledge of SOSs".

When comparing the intentions of city and country schools the results were remarkably even. These results are shown in table 10: Comparison between location and the proposed implementation of SOS for non-pilot schools.

Table 10
Comparison between location and the proposed implementation of SOS for non-pilot schools.

Location	City	Country
Already in use	5	6
1998	4	5
1999	1	2
Has not been decided	16	17

No pattern was determined in the comparison between the school size and the proposed time of implementation of the SOS in mathematics as shown in table 11: Comparison between school size and proposed implementation of SOS in mathematics for non-pilot schools.

Table 11
Comparison between school size and proposed implementation of SOS in mathematics for non-pilot schools.

Number of students	0-99	100-199	200-299	300-399	400-499	500-599	600-699	700-799	800-899
Already in use	3	3	1	2	2	0	0	0	0
1998	1	2	1	1	3	0	0	0	1
1999	0	2	0	1	0	0	0	0	0
not decided	5	1	9	7	5	4	0	1	1

Of the Math pilot schools six of the seven schools who had said they were still using the Math outcomes confirmed this in question 4. One school acknowledged that the trial was only with the Measurement strand and they would be progressing to use the SOS for Mental Computation. The other school, of the seven, said that the implementation would occur in 1998 (similar reasons as given for the non-pilot schools might be inferred from this). The pilot school that was no longer using the Math SOS said that the timing of the implementation had not been decided.

Staff awareness of accountability documents

All schools acknowledged the existence of both the *Improving and Reporting Schools' Performance* (EDWA, 1996a) document and its successor *School Reporting: A Framework for Improving and Reporting* (EDWA, 1997b) in their school.

Of the 56 non-pilot schools 8 (14.3%) said that they had not discussed these documents at a staff level. One of these schools commented that no detailed discussion had occurred. Another asked: "When? We have mostly new staff who are more concerned with the "core business" of teaching. This takes up all their time".

Some of the other 48 (85.7%) schools who answered that discussion had taken place gave comments suggesting that varying levels of discussion and usage of these documents had occurred. Comments ranged from "We have introduced documents to the staff. They are aware of contents", "Cursory discussion only" and "Staff performed this last year" to "Have been ongoing for 2-3 years". One school explained: "As we are a new school the staff are still developing a comprehensive MIS process to justify judgements". Four schools expressed that the union ban had "stifled discussion" two of these adding: "Used by principal and Superintendent to rate the school during quality assurance visits - staff informed" and "Executive team using documents to report to District Office".

One school described how the document was being used: "Teachers are surveyed each term in one area of school performance (eg. Term 3, 1997 - Managing Staff) the results collated and suggestions for 'How we could improve?' acted upon". Another school offered the opinions:

(1) This document has very little impact on what we do as it is only one factor in a range of existing improvement processes.

(2) The manner of 1997 survey of schools means results will be unreliable.

(3) In Fremantle District ¾ items will be self assessed with no real comparability, and ¼ assessed by office.

Of the 8 pilot schools 7 said that there had been discussion of the documents on a staff wide basis. The school that responded in the negative was a school that already had adopted the use of the SOS in Mathematics and had 6 other area SOS being used throughout the school, they commented: "Yes as an overview, but not in detail". Of the seven schools that responded in the affirmative to this question the following comments were made: "But there is very little time available for this", "At a district wide meeting" and "We have been 'through' 2 quality assurance visits and have been rated as excelling in all areas". These comments again show a wide range of discussion, acceptance and usage, as was the case with the non-pilot schools.

Maths Specialists in Schools

Only 10 of the 56 non-pilot schools acknowledged the presence of a Maths Specialist in their school, interestingly, 9 of these were in schools situated in the country. Two of these schools commented that their specialist was in the secondary component of a District High School. In fact, 6 of these 9 schools were District High Schools all having classes from Pre-primary through to Year 10. It was found that a total of 13 District High Schools were represented in the survey and that less than half (46.2%) of

these had a Maths Specialist present. One of the District High Schools that didn't have a Maths Specialist commented that they had a "Science specialist with a strong maths background". Of the 17 Primary schools situated in the country 3 (17.6%) had a maths specialist present. Only one of the 26 (3.8%) Primary schools situated in the city had a maths specialist. One school commented: "It is very difficult for primary teachers to do all this reporting, accountability etc. we have so many areas to teach. High schools have specialists. This must make life a lot less complicated".

Five schools mentioned the presence of a Maths Coordinator or a Key Teacher whose position is to act as a Specialist in this area. The role of these teachers, in two schools was explained as follows: "Maths coordinator with an interest in maths updates staff and attends network meetings and Professional Development where possible!" and "Our Key Teacher ensures teachers are assisted in implementing the Maths area, have appropriate resources, runs a lunchtime maths club for the kids and facilitates the maths review sessions with staff at staff meetings". Two of these five schools have implemented SOS in mathematics.

When compared with the use of the SOS it was shown that of the schools with a maths specialist, none had implemented the use of SOS throughout the school although in two of the schools individual teachers had elected to use the SOS in maths. All 17 schools using SOS in maths had no teacher with specialist training in the area as shown in table 12: Use of SOS compared with the presence of a specialist teacher for non-pilot schools.

Table 12
Use of SOS compared with the presence of a specialist teacher for non-pilot schools.

		Using SOS	**Some teachers using SOS**	**SOS not in use**
Specialist Present	**YES**	0	2	8
	NO	17	2	27

The comparison between the size of the school and the presence of a specialist showed that it was more likely for a specialist to be present in a smaller school, as shown in table 13: Comparison between school size and the presence of a specialist maths teacher for non-pilot schools.

Table 13
Comparison between school size and the presence of a specialist maths teacher for non-pilot schools.

Number of students		0-99	100-199	200-299	300-399	400-499	500-599	600-699	700-799	800-899
Specialist	Yes	1	2	4	1	0	0	2	0	0
Present	No	8	6	7	10	10	1	2	0	2

Of the 8 maths SOS pilot schools only one had specialists in mathematics this school having students from Years 7 through to Year 10, this was also the largest of the pilot schools. Of the 7 other pilot schools, that did not have a maths specialist, 3 were District High Schools all with students from Pre-primary to Year 10+ (one has students in Year 11 and the other two have Year 12 students).

One of the 4 primary schools, none of which had a specialist in mathematics, said that they had a "maths coordinator in charge of a committee and budget area".

Approaches to Accountability in Mathematics

One school (non-SOS) did not answer the question concerning its approach to accountability. The results for the other 55 schools are presented in table 14: How school accountability is being approached in non-pilot schools.

Table 14
How school accountability for mathematics education is being approached in non-pilot schools.

Chosen methods	Number of schools	Percentage
Teacher assessment	38	69.1
Unified school system	18	32.7
Standardised tests	38	69.1
Profile based on outcomes	7	12.7

Of these 55 schools 10 (18.1%) are using teacher assessments only, 3 (5.5%) schools are using standardised testing only and 2 (3.6%) are using profiles only. One school stated that they were using a unified system of assessment only, this was in the form of "open-ended tasks which are moderated by the whole staff and assessed using SOS pointers".

Of the 7 schools that are using profiles one school was using standardised testing as well, whilst three schools were using both standardised testing and a unified system of assessment which unfortunately was not elaborated upon. Three schools which are not using profiles at present stated their intention to do so in the near future, one of these stating that they "will focus on Working Mathematically as a starting point".

One school did not respond positively to any of the methods suggested in the survey but responded that they were using the Education Department's MSE (Monitoring Standards in Education) testing in mathematics in years 3 and 7. This could be seen as a form of standardised testing. Four other schools mentioned the use of MSE testing to help validate the schools results. One school thought that there "should be a state wide test for all students. Not the hit and miss approach that is going on now".

Twenty five schools (45.5%) had chosen to use a combination of teacher assessment and standardised testing. Of these schools 18 (52.9%) were not using SOS in maths, 2 (50%) had at least one but not all teachers using SOS and 5 (29.4%) of the schools were using SOS. This shows a possible trend away from this combination of assessment procedures with the use of SOS in mathematics.

As well as MSE as a form of standardised testing the Australian Schools Mathematics Competition that is run annually by the University of New South Wales was mentioned, as was Easy-mark. (Easy-mark is a company that will provide the school with standardised tests which are administered and returned to the company for marking and grading. The school would have to allow for the cost of this in their budget). One comment stated "Easy-mark used to supplement SOS (unified system) judgements. The Easy-mark standard Space, Measurement and Number tests are very 'Wood and Lowther' posed and we're not quite ready for the 'outcomes' Easy-mark maths test".

Two schools mention that they use the data from the school's reports to give a picture of the school performance in maths. One school added: "Performance can be monitored in successive calendar years or track a class/year result through the progress from Yr 1 - 7".

A comparison of methods of accountability compared with the school's usage of SOS is presented in table 15 : Comparison of methods of accountability used with the school's use of SOS, in non-pilot schools.

Table 15
Comparison of methods of accountability used with the school's use of SOS, in non-pilot schools.

Methods of assessment used		Using SOS		Some teachers using SOS		SOS not in use	
		no.	%	no.	%	no.	%
Teacher assessment	Yes	10	58.8	2	50	27	79.4
	No	7	41.2	2	50	7	20.6
Unified school system	Yes	9	52.9	2	50	7	20.6
	No	8	47.1	2	50	27	79.4
Standardised tests	Yes	10	58.8	4	100	24	70.6
	No	7	41.2	0	0	10	29.4
Profile based on outcomes	Yes	6	35.3	1	25	0	0
	No	11	64.7	3	75	34	100

NB The percentages in the table are of the SOS grouping and not the total population.

As can be seen from the table the use of teacher assessment for accountability purposes tends to decrease with the use of the SOS, whilst the use of unified systems of assessment tend to increase. Although standardised testing is popular in all groups the data shows a slight lower use of standardised tests where the SOS are being used. The group using SOS shows a higher use of student profiles.

Table 16
Comparison of methods of accountability used with the school's location, in non-pilot schools.

Methods of assessment used		City schools		Country schools	
		no.	%	no.	%
Teacher assessment	Yes	23	88.5	15	51.7
	No	3	11.5	14	48.3
Unified school system	Yes	5	19.2	13	44.8
	No	21	80.8	16	55.2
Standardised tests	Yes	17	65.4	21	72.4
	No	9	34.6	8	27.6
Profile based on outcomes	Yes	2	7.7	5	17.2
	No	24	92.3	24	82.8

NB The percentages in the table are of the city/country grouping that responded to the question and not the total population

From the above table, table 16: Comparison of methods of accountability used with the school's location, in non-pilot schools, it can be seen that schools in city areas rely more on teacher evaluation than do the schools in country areas for their accountability. A higher percentage of the country schools surveyed have developed a unified system of assessment than had the city schools. Standardised testing is prevalent in both areas whereas the use of profiles based on the outcome statements is not.

Of the 17 schools using SOS 8 (47.1%) were in the city and 9 (52.9%) were located in the country. Of the schools using outcome statements a greater percentage of country schools were using profiles than were the city schools. (ie. 5 out of 9 country schools and only 2 out of 8 city schools).

No noticeable pattern emerged when the method of assessment was compared with the size of the school.

For the pilot schools, table 17: How school accountability for mathematics education is being approached in the pilot schools, shows a similar pattern to table 14 where the accountability of non-pilot schools was shown. A unified school system and a profile based on outcomes are used on fewer occasions than are teacher assessments and standardised tests.

Table 17
How school accountability for mathematics education is being approached in the pilot schools.

Chosen methods	**Number of schools**	**Percentage**
Teacher assessment	4	50
Unified school system	3	37.5
Standardised tests	4	50
Profile based on outcomes	1	12.5

The school that is using a profile based on SOS said that teacher assessment "still occurs" as well as the use of standardised testing, however, a unified system of assessment is being developed using the outcome statements in the sub-strands of measurement and mental computation this will form the basis of their profile.

Another school that uses teacher assessments and standardised testing stated: "Standardised testing programme is still under trial and profile being established".

One school only marked the unified system of assessment as being used, however, in the further comment section said that they were using MSE, Easy-mark, SOS tests and teacher tests.

Another school which had indicated that each teacher assessed their own programme in their own way went on to remark: "Each teacher assesses their own program however it is based upon common tasks which are moderated".

No mark was placed on one survey but the school did comment: "Each team of teachers (5 or 6) collaborates to design an appropriate method of monitoring student outcomes in maths. There is also an overall school monitoring system required by Maths coordinator (according to plan)".

The school that is no longer using the SOS indicated that the teachers in the school assessed their own program and that standardised testing is used, two other schools also marked this combination although one school also marked that they were developing a unified system of assessment.

Comparing the group of pilot schools which are still using SOS with the group of non-pilot schools who have adopted the use of SOS in maths, a slightly lower percentage of pilot schools used teacher assessment (42.8% as compared to 58.8%), a unified approach (42.8% as compared with 52.9%), standardised testing (42.8% as compared to 58.8%) and a profile based on SOS (14.3% compared to 35.3%).

Table 18
Comparison of methods of accountability used with the school's location for pilot schools.

Methods of assessment used		City schools no.	City schools %	Country schools no.	Country schools %
Teacher assessment	Yes	1	25	3	75
	No	3	75	1	25
Unified school system	Yes	2	50	1	25
	No	2	50	3	75
Standardised tests	Yes	2	50	2	50
	No	2	50	2	50
Profile based on outcomes	Yes	1	25	0	0
	No	3	75	4	100

NB The percentages in the table are of the city/country grouping that responded to the question and not the total population

Table 18: Comparison of methods of accountability with the school's location, looks at how the pilot schools situated in the city and country areas responded to the methods of accountability used. This does not follow the same pattern as the non-pilot schools where 88.5% of the city population and 51.7% of the country schools were using teacher testing, here there is a lower percentage of city and a higher percentage of country schools.

With the development of a unified system of assessment there is a higher percentage of city schools (50% as compared to 19.2% non-pilot) and a lower percentage of country schools (25% as compared to 44.8% non-pilot) using this method of assessment.

The use of standardised testing is slightly lower for both the city (50% as compared to 65.4% non-pilot) and country (50% as compared to 72.4% non-pilot) pilot schools.

The use of profiles was low for both city and country pilot schools as it was for the non-pilot schools.

Preferred types of assessment

One of the non-pilot schools did not respond to the question concerning the types of assessment used to monitor the students' development. The results for non-pilot schools are shown below in table 19.

Table 19
Forms of assessment used to monitor students' development in mathematics education in non-pilot schools.

Assessment type	No. of schools	Percentage
Teacher tests	54	98.2
Checklists	39	70.9
Portfolios/Work samples	31	56.4
Anecdotal records	36	65.5
Student journals	7	12.7
Standardised tests	43	78.2
Student self-assessments	13	23.6
Principal's tests	1	1.8

It can be seen that teacher tests are the most common form of assessment of mathematical development in the surveyed schools, being used by all but one school, interestingly this school is one of those that have adopted the use of the SOS.

The use of standardised tests is also very common. The standardised test is used in more schools for the assessment of student development (78.2% of schools) than for accountability purposes (69.1%).

Checklists and anecdotal records are quite widely used, with portfolios of work samples being used by just over half the schools.

Student self-assessments and journals are used by only a few schools whereas the principal's test is rarely used.

Other forms of assessment that were suggested by schools include the analysis of open ended tasks, whole of year tests, The Australian Mathematics Competition (which could be included as a specific standardised test) and Principal's interviews (which has been introduced by one school since accountability pressures). One school mentioned that Easy-mark had been used for diagnostic purposes.

Table 20
Comparison between forms of assessment used to monitor student development and the use of SOS in non-pilot schools.

Types of assessment used	Using SOS		Some teachers using SOS		SOS not in use	
	no.	%	no.	%	no.	%
Teacher tests	16	94.1	4	100.0	34	100.0
Checklists	12	70.6	3	75.0	24	70.6
Portfolios/Work samples	13	76.5	3	75.0	15	44.1
Anecdotal records	11	64.7	3	75.0	22	64.7
Student journals	3	17.6	0	0	4	11.8
Standardised tests	12	70.6	4	100.0	27	79.4
Student self-assessments	5	29.4	2	50.0	6	17.6
Principal's tests	0	0	0	0	1	2.9

NB The percentages in the table are of the SOS grouping and not the total population.

Table 20: Comparison between forms of assessment used to monitor student development and the use of SOS in non-pilot schools, looked to see if there were differences between the school's level of adoption of SOS and their preferred methods of assessing the students. The use of each of these forms of assessment is fairly well balanced amongst the varying forms of acceptance of the student outcome statements. The only case where there is a noticeable difference is in the use of portfolios containing samples of students' work. These are used more by the schools that have adopted SOS than by the ones that have not.

Apart from the use of anecdotal records, there was very little difference in the use of these assessment strategies between the city and the country areas (76.9% of city schools and 55.2% of country schools were using anecdotal records).

When the size of the school was compared with methods of assessment used it was found that the use of teacher tests, checklists, standardised tests and principal's tests was fairly consistent for all school sizes. Portfolios and anecdotal records were mainly used in the schools from 100 to 499 students in size, whilst student journals were only used in schools with 100 to 399 students. Student self-assessments had a very low percentage of schools using them except for in schools with 200 to 299 students where 54.5% of the schools were using them.

Most schools indicated that they were using several forms of assessment. Only two schools indicated one assessment and both of these were using teacher tests. Six schools marked two forms of assessment, half of these were using a combination of teacher tests and standardised tests, one school was using teacher tests and checklists, another was using teacher tests and portfolios and the other was using anecdotal records and portfolios. Eleven schools indicated three forms of assessment with seven of these being teacher tests and standardised tests with either anecdotal records or checklists. Thirteen schools were using four forms of assessment, 15 were using five, 6 using six and 2 using seven (these only omitted the principal's test).

Table 21
Forms of assessment that are used to monitor the students' development in mathematics education in pilot schools.

Assessment type	No. of schools	Percentage
Teacher tests	8	100
Checklists	7	87.5
Portfolios/Work samples	7	87.5
Anecdotal records	6	75.0
Student journals	6	75.0
Standardised tests	5	62.5
Student self-assessments	5	62.5
Principal's tests	1	12.5

The types of assessment used by non-pilot schools to monitor student progress is shown in table 21: Forms of assessment that are used to monitor the students' development in mathematics education in pilot schools. In addition to this, one school mentioned the use of open-ended integrated tasks as a form of assessment used to monitor students' development.

The school that is using the principal's test is the pilot school that is no longer using any SOS.

There were no noticeable patterns when the responses of the pilot schools were compared with the schools size or location.

One school said that they were using all 8 forms of assessment, 2 schools were using seven (both omitting principal's tests), 1 was using six, 2 were using five, 1 was using four and 1 was using three of the assessment types.

A considerably higher percentage of the pilot schools are using student journals and self assessments (75% and 62.5% respectively) than the non-pilot schools (12.7% and 23.6% respectively) or the non-pilot SOS schools (17.6% and 29.4%) and a

slightly smaller percentage are using standardised testing (62.5% for pilot schools to 78.2% non-pilot and 70.6% non-pilot SOS schools).

On average the pilot schools (5.6 types of assessment) were using more assessments than the non-pilot schools (4.1) and the non-pilot schools that are using SOS (4.2), however the pilot school that said they were no longer using SOS said they were using all 8 forms of assessment (this seems to be something of an anomaly).

Assessments introduced due to either SOS or accountability measures

Three non-pilot schools did not respond to the question concerning the types of assessments that had been introduced as a result of either the introduction of SOS or accountability measures, two of these were from the SOS group the other was from the non-SOS group. The results for both non-pilot and pilot schools are shown in table 22: Assessments introduced as a result of SOS or accountability in non-pilot and pilot schools.

Table 22
Assessments introduced as a result of SOS or accountability in non-pilot and pilot schools.

Assessment type	No. of non-pilot schools	Percentage non-pilot schools	No. of pilot schools	Percentage pilot schools
Teacher tests	8	15.1	0	0
Checklists	8	15.1	1	12.5
Portfolios/Work samples	17	32.1	2	25.0
Anecdotal records	8	15.1	0	0
Student journals	3	5.7	3	37.5
Standardised tests	14	26.4	1	12.5
Student self-assessments	2	3.8	0	0
Principal's tests	0	0	0	0

In the non-pilot schools the most noticeable introductions are of the portfolios and the standardised tests. With the portfolios 8 (53.3% of the group) were from the group that have introduced SOS, 2 (50% of the group) from the group that have some teachers using SOS and 7 (21.2%) from the non-SOS group, showing that a higher percentage of schools using SOS had introduced the use of portfolios. With standardised tests 5 (33.3%) were from the SOS group and the other 9 (27.2%) were from the non-SOS group. The two schools that had introduced student self-assessments were one each from the some SOS and the non-SOS groups. Two of the three that had introduced student journals were in the SOS group and the other was in the non-SOS group.

Twenty seven (51.9%) schools indicated that there had been no new assessment techniques introduced as a result of either of the documents. Eleven schools had introduced just one form of assessment six of these introducing standardised testing, four introduced portfolios and one introduced anecdotal records. One school indicated that they had introduced seven types of assessment ie. everything but the principal's test (this causes one to ask what they had been doing before? - or had the question been misinterpreted?).

One school mentioned the introduction of open-ended tasks "which are moderated by the whole staff and assessed using SOS pointers".

In the pilot schools the most noticeable introduction was in the use of student journals. Three schools indicated that they had introduced no new methods of assessment, one school commenting that they "have developed the ones mentioned previously as our understandings increase".

Opinions concerning which assessments do not give valid and reliable information.

Five schools from the non-pilot group did not respond to the question asking for an opinion of which assessments were considered to lack validity and reliability. The results for both non-pilot schools and pilot schools are presented in table 23: Opinions concerning the lack of validity and reliabilty of assessment types.

Table 23
Opinions concerning the lack of validity and reliability of assessment types.

Assessment type	No. of non-pilot schools	Percentage non-pilot schools	No. of pilot schools	Percentage pilot schools
Teacher tests	4	7.8%	0	0%
Checklists	4	7.8%	1	12.5%
Portfolios/Work samples	3	5.9%	0	0%
Anecdotal records	4	7.8%	1	12.5%
Student journals	4	7.8%	0	0%
Standardised tests	8	15.6%	1	12.5%
Student self-assessments	8	15.6%	2	25.0%
Principal's tests	15	29.4%	2	25.0%

Of the non-pilot schools who responded twenty eight schools (54.9%) indicated that they thought all forms of assessment gave valid and reliable information with comments such as: "All provide some insight" and "We believe all/any form of assessment contributes to 'the big picture' and is therefore valid/reliable". Some thought that it depended on the individual assessment: "All of the above can be excellent or useless depending on the quality and suitability of the materials". One thought that it depended on the purpose of the assessment: "The value of each form is related to how the information is to be used eg. a student journal or portfolio is of use in showing progress and in parent discussions whereas a checklist of standardised test is little use in that domain". Two schools stated that several forms of assessment

should be used in conjunction to validate each other: "On their own, none!" and " No one form of assessment is reliable. However, several forms used in conjunction are more likely to give more reliable information". In contrast to these, one school was more cautious: "All forms have to be treated with a great deal of care".

Two schools commented on bias in standardised tests: "Standardised tests depend on the state it has been initiated from - can be culturally biased" and "We have 65% Aboriginal students most of which have reading problems. There is a cultural bias in most tests". Two schools commented on the Principal's test: "May not be relevant unless the Principal knows what he/she is doing" and "Principal's tests are a thing of the past and I don't know of a school that would conduct them".

When the use of SOS was used to compare responses within the non-pilot group it was found that the only differences worth noting are with respect to standardised testing and self-assessments. Five (33.3%) of those using SOS questioned the reliability and validity of standardised tests as compared to 3 (9.4%) of the non-SOS group. Only one (6.7%) of those using SOS questioned the reliability and validity of student self-assessments, one (25%) of the schools with some teachers using SOS and 6 (18.7%) of the non-SOS group did likewise.

There were no noticeable differences when compared according to location or size of school.

Three of the pilot schools thought that all forms of assessment were reliable:

> All give useful data. Getting an accurate picture of a child's maths performance is like looking at an oil painting. Any one small part may be messy, uni-coloured have ugly brush marks and seem meaningless. When you put enough different parts

> together and stand back a bit the picture becomes obvious, meaningful etc. When you've collected enough different bits of data on student performance the child's overall performance level looks fairly obvious!

One of the two schools that marked student self-assessments as not giving valid and reliable data clarified their decision by adding: "Student self-assessment can be a useful tool - in some areas and in addition to other forms of evaluation".

A school that marked Principal's tests asked: "What are Principal's tests?"

People involved in the school's response to accountability.

Question 11 in the survey asked for an indication of which parties within the school had input into the decisions about how the school had approached accountability. The results are presented in table 24: People involved in deciding the school's response to accountability in non-pilot and pilot schools.

Table 24
People involved in deciding the school's response to accountability in non-pilot and pilot schools.

	No. of non-pilot schools	**Percentage non-pilot schools**	**No. of pilot schools**	**Percentage pilot schools**
The Principal	52	92.9	8	100
Some teaching staff (only)	7	12.5	0	0
Parent groups	28	50.0	5	62.5
Deputy Principals	35	62.5	6	75.0
All teaching staff	48	85.7	7	87.5

The four schools, that did not include the principal, had marked that all teaching staff had participated and therefore the principal is probably included (ie. the schools had

interpreted the question differently). One school marked the principal only, adding that there are no deputies. This school is in its first year and so all of its systems are in the early stages of development. One other school commented that there are no deputies, in fact all 9 schools under 100 students in size and 6 of the eight schools between 100 and 199 students in size did not include deputies possibly because they do not have them. Of the schools with over 200 students only 6 schools did not include the deputies, 5 of these had marked 'all teaching staff' therefore possibly including the deputies in this response the other one is the above mentioned school in its first year.

Whether the school is using SOS and the location of the school seem to have no effect on the responses given.

One school clarified its parent involvement by adding that: "SBDMG (School Based Decision - Making Group - which has staff as well as parent representation) reps are present during annual MIS review".

The statistics for the pilot schools are fairly similar. One school only did not have full teacher support and most of the schools had parent involvement. Three schools mentioned the involvement of the school council or SBDMG. One school commented "Parent groups not formally consulted prior to changes but provide feedback".

The effect of accountability on staff communication

Question 12 of the survey suggested that the *Improving and Reporting Schools' Performance* document had increases the communication between the staff members and asked for the principals to comment on this statement from 'strongly agree' to 'strongly disagree'. The results for non-pilot and pilot schools are presented in table

25: How accountability has affected the communication between teaching staff in regard to assessment of mathematics.

Table 25
How accountability has affected the communication between teaching staff in regard to assessment of mathematics.

	No. of non-pilot schools	**Percentage non-pilot schools**	**No. of pilot schools**	**Percentage pilot schools**
Strongly agree	5	8.9	1	12.5
Agree	8	14.3	4	50.0
Undecided	16	28.6	2	25.0
Disagree	21	37.5	0	0
Strongly disagree	3	5.4	1	12.5
No response	3	5.4	0	0

The three non-pilot schools which did not register a ticked response each wrote a comment of explanation: "Not in use due to union ban", "Made no difference. Its just an endless stream of EDWA policies" and "May in future when used".

One of the schools that strongly disagreed with the statement wrote that they had had a "cursory discussion only".

Some of the schools that disagreed also made comments. Two of these also attributed the lack of discussion to the union ban. The school that is in its inaugural year stated: "It should do but not here yet". A final comment explained how that school handled its own approach to accountability: "Our assessing regime flows from reports data and feeds our management information system of our performance indicators".

Of the non-pilot schools using SOS that responded to this question 12.5% strongly agreed, 25% agreed, 31.25% were undecided and 31.25% disagreed whilst the schools where some of the teaching staff were using SOS responded with 33.3%

agreeing and 66.7% being undecided and of the non-SOS schools 8.8% strongly agreed, 8.8% agreed, 32.6% were undecided, 41.2% disagreed and 8.8% strongly disagreed.

Of the pilot schools the undecided school commented that "these were already happening". Two of the schools that agreed with the statement wrote comments: "Opened discussion" and "highlighted the need for moderation". The school that strongly agreed said that this was for "all areas though".

In comparison the pilot schools have a higher modal score ('agree' as compared to 'disagree').

When given a numerical value of 5 - strongly agree, 4 - agree, 3 - undecided, 2 - disagree and 1 - strongly disagree the mean and standard deviation for the non-pilot group were 2.83 and 1.05 respectively, and for the pilot group they were 3.5 and 1.11 respectively. The mean for the pilot group was higher and above three showing that they were more inclined to agree with the statement. The standard deviations were similar showing that there was no difference in the spread of results between the two samples.

The effect of accountability on teacher/community communication

Question 13 of the survey suggested that the *Improving and Reporting Schools' Performance* document had increases the communication between the teaching staff and the wider school community and asked for the principals to comment on this statement from 'strongly agree' to 'strongly disagree'. The results for non-pilot and pilot schools are presented in table 26: How accountability has affected the communication between teaching staff and wider school community in regard to assessment of mathematics.

Table 26
How accountability has affected the communication between teaching staff and wider school community in regard to assessment of mathematics.

	No. of non-pilot schools	**Percentage non-pilot schools**	**No. of pilot schools**	**Percentage pilot schools**
Strongly agree	4	7.1	1	12.5
Agree	9	16.1	1	12.5
Undecided	11	19.6	4	50.0
Disagree	26	46.4	1	12.5
Strongly disagree	3	5.4	1	12.5
No response	3	5.4	0	0

The comments from the no response group were the same as for the previous question.

A school that strongly disagreed commented: "We have published invitation to all parents (ie. over 1000) to view our MIS results of June '95, Dec '95, June '96, Dec '96, June '97 and also 100 family survey of parent and school and not 1 person has taken opportunity!!!"

Of the schools that agreed two wrote: "Not just in math" and "Ed council reporting processes".

One of the schools that strongly agreed added: "There has always been strong consultation and participative decision making between school/parents".

When compared using the SOS groupings for the non-pilot schools the percentages are as follows: SOS group 12.5% strongly agree, 37.5% agree, 18.75% undecided, 31.25% disagree; some teachers using maths SOS group 66.7% undecided and 33.3% disagree, and the non-SOS group 5.9% strongly agree, 8.8% agree, 17.6% undecided, 58.8% disagree and 8.8% strongly disagree.

For the non-pilot schools the modal score was 'disagree' whilst in the math pilot schools the modal score was 'undecided'.

When the results were allocated a numerical value (5 - strongly agree.....1 - strongly disagree) the mean for the pilot schools was marginally higher than that of the non-pilot schools (3 compared with 2.72), the standard deviations were also similar (1.12 for pilot and 1.05 for non-pilot schools).

School to community communication methods

Question 14 of the survey asked the schools to indicate the methods that had been used by the school to inform the wider school community of the developments in the assessment and accountability in mathematics. Five schools from the non-pilot group did not answer this question four of which were from the non-SOS group. The results for both the non-pilot and pilot schools are presented in table 27: Methods used to inform the wider school community of the developments in assessment of mathematics.

Table 27
Methods used to inform wider school community of the developments in assessment of mathematics.

	No. of non-pilot schools	**Percentage non-pilot schools**	**No. of pilot schools**	**Percentage pilot schools**
School news letter	31	60.7	5	62.5
Student reports	33	64.7	7	87.5
The P&C	28	54.9	6	75.0
Special letters	6	11.8	2	25.0
Parent Evenings	18	35.3	4	50.0

In addition to this fourteen of the non-pilot schools and one of the pilot schools mentioned the School Council or School-based Decision Making Group, three schools

referred to the SBDMG's annual report. One of the non-pilot schools said that they had "school open days" whilst another described their "maths activity days where parents are invited to come and work with their child and a group of children". Two schools said that no methods were used and another commented that "one on one contact" was the only method used.

One of the pilot schools mentioned the use of "work sample files".

Within the non-pilot group it was found that the only real difference between the SOS and non-SOS group was in the use of parent evenings where 58.8% of SOS schools and 11.8% of non-SOS schools said that they used them.

In comparing the pilot schools with the non-pilot schools it was found that on average the pilot schools used more methods (3) of informing the wider community than did the non-pilot schools (2.3).

Reports on the schools which were followed up.

School 1.

The first school became a pilot school in 1995 when it embarked on the introduction of the Measurement Strand of Student Outcome Statements throughout the school. Mathematics had been earmarked as one of the School Priority Areas in need of development. When the staff were invited (as were all other schools) to join the pilot programme, and as the school had already identified the need in this area, they accepted the offer.

A lot of assistance was given to the staff of the school in the form of modelled lessons with subsequent moderation meetings where groups of staff members looked at the work produced by the children and tried to determine which level they fitted into. At first there was much disagreement about the levels that the students' work

indicated and much discussion was entered into. As the staff became more confident in the "levelling" of students there were less discrepancies and more agreement and consistency in the decisions made.

Each child in the school has a progress card to indicate the levels that they have achieved. The following is an example of the progress card used.

Table 28
Example of student progress record card.

RECORD OF MATHEMATICS PROGRESS

STRAND	SUBSTRAND	LEVEL									
		1		2		3		4		5	
MEASUREMENT	CHOOSE UNITS										
	MEASURE										
	ESTIMATE										
	TIME										
	USE FORMULAS										

A single tick is given when a student is substantially achieving a given level, two ticks are given when the student has definitely attained the level.

According to the school's acting Principal one advantage of this system was realised when the school moved into its second year of using the Student Outcome Statements. He commented that the continuity of the system as the child progresses from year to year is seen as a definite benefit and is the real strength of the programme. Another benefit that the school has noticed is that as the area of measurement was given the status of a school priority and all staff have participated enthusiastically the result has been an improvement in the abilities of the students in this strand. This improvement has not transferred to other areas of mathematics. This could be attributed to the change of focus that the staff have experienced in their

teaching from a system concerned with what the teacher wants to teach next and assessing what the children can and can't do, to a system that is concerned with where the students are and what they need to be able to do next. The acting Principal commented that this produced more efficient teaching, the teachers are teaching more thoroughly and assessing more thoroughly.

The assessment of the students has occurred in several ways including performance tasks that give rise to anecdotal records, work samples that are collected and stored in scrap books and check lists. Assessment is seen as ongoing and observations are made over a period of time. Standardised tests, in the form of the Primary Schools Mathematics Competition which is conducted annually by the University of New South Wales, has been used not to assess the students but to substantiate and justify the school's results for accountability purposes. The children at this school seem to be below average in their development and the standardised test supports this. There is concern in the school, about the use of standardised tests. Especially the language usage in tests cause 'heavy reservations'.

For the purpose of reporting the teachers convert the student's achieved levels back to the old categories of 'excellent progress', 'satisfactory progress', 'below standard' and 'well below standard'. The school sees a problem with reporting to parents in levels and believes that the parents will need to be educated in the concept behind the Student Outcome levels before the system of reporting in levels will be accepted. A parent would be upset if their child did not progress through the levels in sequential reports. It is quite conceivable that a student might stay at the one level for a period of two years or more, this would appear to be unacceptable to parents unless they had been informed that this situation is quite normal. The school has decided not to embark on this step until they have seen the reporting information being produced

by the Curriculum Corporation as part of the Curriculum Framework package. The school's Curriculum Framework coordinators will familiarise themselves with the reporting plan next year and they will in turn educate the rest of the staff.

In 1998 the school will extend its use of the Student Outcome Statements in Mathematics to encompass the Mental Computation substrand of the Number strand. This has been put forward by the teaching staff to the School-based Decision Making Group and has been presented in the school plan to the superintendent for ratification.

The one disadvantage that the staff can see with using Student Outcome Statements is that there are few other schools doing the same thing at the present time. Once the use of the Outcome Statements is system wide, they think that it will be beneficial to compare the breakdown of levels between schools, although they feel that there will still need to be substantiation of these levels with student performances on a standardised test.

With the target of every school having the Curriculum Framework in full operation in all areas of the curriculum by the year 2003 the school feels that they are already on track to meet these requirements.

School 2

This school is situated in an inner city suburb where the school population is 80 - 90% from a non-English speaking background. An extensive ESL (English as a Second Language) programme is operational within the school. Due to the great diversity of backgrounds that the children come from there is a wide range of culturally based learning styles that new students come with. As the School Development Coordinator explained, "the Chinese for instance are very much rote learners, they come with a strong command of basic number facts, they're very good

at number, whereas from an Eastern European or African background the number concepts are not as good as the Chinese they're not visual learners, they're more hands-on learners - very concrete". Mathematics was identified in 1994 as a priority area not only because of the diverse nature of the students but also because it was as an area where the teachers at the school could further develop their teaching skills. For these reasons the school became part of the First Steps programme commencing in 1994 with a view to it being a three year programme.

In 1995, as part of the monitoring of the First Steps programme within the school one of the open-ended tasks, provided by the programme, was selected to be administered school wide. The first task to be selected was called Drawing Shapes. As a staff the teachers discussed how the task would be carried out. Slight changes were made to cater for the Junior Primary children in the school. The Middle and Upper primary teachers decided to extend the activity. The task was initially administered in March. One of the school's Mathematics Coordinators was present in the classroom to assist with the moderation of the task being conducted, making sure that each teacher explained the task sufficiently but without giving the children too much assistance. The presence of this additional person in the classroom was also beneficial for the purpose of observation. Whilst the classroom teacher was preoccupied with individual students as they asked questions, this extra person could listen, observe and make anecdotal records of what the students were doing.

Two aspects of the task, 'skills in construction' and 'skills in shape', were selected to be analysed from the students' work samples and the anecdotal records that had been collected. The teachers then used the document *Making the Links* (EDWA, 1995) to decide where the individual students were placed in relation to the Student Outcome Statements.

In *Making the Links* (EDWA, 1995) the Student Outcome Statements are placed across the top of the page in progressive levels, at the bottom of the page are the stages as outlined by the *Learning Mathematics* (EDWA, 1989) syllabus which roughly correspond to the students' year levels. The two continuums are synchronised so that a vertical movement of the eyes gives the reader a rough comparison between the Student Outcome Statements and the stage that the student can be expected to be working at. For example, in relation to the construction of shapes a student who is in year four, if working at stage four, could be expected to have completed level 2 (SOS) and be working toward level 3 (SOS). If the student, from samples of work and observation, is identified as performing at level 1 (SOS) then they are not performing at the stage of the syllabus corresponding to their year at school.

The teachers collaborated to decide the levels of the Student Outcome Statements that the individual students were working at, according to the pointers given in *Making the Links* (EDWA, 1995). For some students this was difficult as they were observed to copy from a neighbour or as no attempt was made at all. The student's level was then translated to the format used on the school's report card. The school's reporting continuum consists of the headings: 'Consistently', 'Usually', 'Developing' and 'Cause for concern'. 'Consistently' would indicate that the student was always working at or above the level of Student Outcomes that relate to student's year level. 'Usually' would indicate that the student was showing considerable indication of being at an appropriate level of SOSs to the year at school. 'Developing' would indicate that the student was showing some indication of performance at the level (SOS) appropriate to the school year level. 'Cause for concern' indicates that the student is failing to show any sign of being at an appropriate level (SOS) and that the student's progress is very slow.

The student's report levels were then collated and graphed to give an overall school picture.

The school staff used the obtained information to plan their teaching programme over the next few months to help their students develop in the chosen sub-strand. The *Learning Mathematics* (EDWA, 1989) syllabus is heavily relied on as a source of developmental activities. Along with this development in the teaching programme, the school's resources in the form of teaching aids and additional curriculum materials were upgraded to support this development and are then available in subsequent years as teachers repeat their innovations. The school's Math Coordinators were available to assist teachers in course developments, making suggestions, helping to implement activities that could be used and monitoring progress within the classroom through the discussion of the children's work samples with the teacher. The Math Coordinators were especially interested in the development of students who were seen to be 'at risk' and they monitored their progress through their work samples.

The work samples were also used for liaising with parents and were sent home on a regular basis.

Toward the end of term 3, that is , in about August or September the task that was implemented in March was re-administered to determine if the intervening activities have been successful. The work samples of the individual students from March and September were compared to see how each student had progressed.

Again, the students' work was analysed and translated into a report score. The overall results were graphed and compared with the previous results with improvements being noted. One such improvement for the 1995 results showed that

although 1.5% of the students had made no attempt in the March assessment no student fitted into this category in the August assessment.

For accountability purposes the District Superintendent was shown the overall results, the school plan for improvement that was followed is shown and the data is supported by the work samples of the children from varying levels.

In 1996 the activity called "The Number Ten" from the First Steps material was used as the assessment tool to develop some number concepts and this year (1997) the measurement task "How Big are You" was used. Next year the school intends to return to the Space strand, but will explore a different sub-strand to the one previously used. All previously used sub-strands are considered to be a part of the teachers' repertoire, although this presents a problem with new staff to the school. It might be noted that the First Steps materials include the SOS strand of Chance and Data as part of the number strand and that the SOS strand Working Mathematically is a focus of all First Steps materials.

The teaching staff have found the *Improving and Reporting Schools' Performance* (EDWA, 1996a) document to be very beneficial. Of special interest and the basis for much discussion has been the areas of 'Teaching and Learning', 'The Learning Environment' and 'Improving Students' Performance'. The School Development Coordinator stated that the document "showed us the place of assessment in the school development plan". An awareness developed, from the discussion about the document, that assessment had to be linked to the performance indicators, it had to show the students' development and improvement. The purpose behind the assessment became apparent. He stated that, "we're not just collecting data for the sake of it". As the current system of assessment was developed before the teachers' awareness of the *Improving and Reporting* (EDWA, 1996a) document it

was decided to investigate how many of the pointers were already being covered. As it was found that most of the pointers were already being addressed it was decided to continue with that system.

Communication with the parent body has also been strong. Parents have been invited to attend School Development Day sessions. The parents were involved with discussions concerning what was currently being achieved, what the desirable indicators of school success would be, what areas of concern there were and the development of future recommendations. The S. D. Coordinator commented, "the parents have been involved in that review process for three years". In general, the parents of the school are very actively involved in the P&C and in the decision making groups although this representation is not from all ethnic groups and is mainly comprised of the parents from English speaking backgrounds. A concerted effort is being made to involve the non-English speaking parents. All parents are welcome to visit the classrooms of their children at any time and special days are also set aside for parents to visit and be shown what their children are doing in the priority areas. Teacher interviews and the school newsletter are other avenues of communication.

Standardised tests are seen as having little value in the school due to their heavy reliance on language and the cultural biases they present.

A portfolio of the student's work is sent home at the end of terms one and three and is seen as more reliable evidence of the student's progress. The coordinator made it clear that they "put in samples of their (the students) true work not just work that's nice but their true everyday work so that the parents can see and can understand where the children are at. With the portfolio there is an explanation, the teachers can write a comment about the level that the children's work is at. The work sample from the end of term one is kept and then they do another one at the end of term three so

that they can visually see the progress they've made. Of course when the children do a piece of work that shows something about what we are doing with the Student Outcome Statements, that is dated and put into the children's file to be taken home at the end of term one or term three". These same portfolios are used to justify the results given to the superintendent.

It was strongly emphasised on several occasions in the interview that progress should be slow and thorough. One small sub-strand, it was reiterated, is sufficient to tackle each year (especially as this is being done in each of the eight learning areas in the school curriculum and not just in Mathematics). It was pointed out that each teacher has to come to terms with the language, which in some places has different meanings, being used in the Student Outcome Statements, as well as the change in emphasis that it is directing, hence the slow and steady nature of the implementation that is occurring. The value of the *Making the Links* (EDWA, 1995) document to the progress in the school was also greatly emphasised.

"Hopefully, by the year 2000 or 2002 we will have some expertise in some areas of the Student Outcome Statements. There'll probably be a lot more to learn at that time but at least we've made a start. We now consider ourselves to be experts in the areas that we have done", the S. D. Coordinator stated.

School 3

One of the three Mathematics Coordinators of the third school said that the school had not embarked on any implementation of the Student Outcome Statements at that time as they were under the impression that only the SOS Pilot Schools had been able to obtain copies of them. He described the "frustration, misunderstanding, angst and confusion" that existed in the school as the teaching staff are well aware of the

immanency of the Student Outcome Statements and yet they have had no feedback from the Education Department about where they are meant to go.

The Student Outcome Statements would begin to be used in some classrooms as soon as they are available. The school foresees that there is much to be learnt and as soon as possible they will begin to "chip away at it" keeping in mind that they have only a few years to have it all in place.

Up to now the school feels that there has been insufficient support and the documents that have been issued such as *Improving and Reporting Schools' Performance* (1996a) have had no positive input in the discussion process within the school or its wider community.

Next year the school hopes to work closely with a local primary mathematics textbook writer as he designs assessment items to support his textbook materials. The school expects to trial the performance activities that he develops.

Discussion.

As the Student Outcome Statements have not been officially introduced in all State schools, the number of schools, from the survey, which have voluntarily begun to use them is very promising. Many of these schools made very positive comments about the SOS and the improvements that they have noticed in their teaching and in the levels of their students through their use. On the other hand, the survey showed that there is a great amount of caution and concern about the introduction of SOS and a strong reluctance to the changes inherent in the use of SOS as indicated by some of the comments from schools that have not as yet decided on their own implementation of them.

Although the documents *Improving and Reporting Schools' Performance* (EDWA, 1996a) and *School Performance: A framework for Improving and Reporting* (EDWA, 1997b) have been received by all of the schools which responded to the survey, a wide range of discussion, acceptance and implementation was noted in both the pilot and non-pilot schools. In several schools the documents had not been discussed with the school teaching staff and were being used for MIS purposes by the Principal only. Where discussion was indicated by the schools, the extent of the discussion varied. In some schools this was indicated to have been of a cursory nature only, whilst other schools indicated extensive discussion and usage of the documents by the staff, in their comments.

The presence of a trained maths specialist seemed to have no effect on whether the school had implemented the SOS in maths. In fact, a strong correlation was seen in the survey data between the fact that a school had a maths specialist and the non-implementation of SOS in maths, as none of the 10 non-pilot schools that acknowledged the presence of a specialist mathematics teacher had implemented SOS on a school wide basis. Of the 33 pilot and non-pilot schools using the SOS in maths only one had the expertise of specialists in the area.

The predominant forms of assessment being used for accountability purposes are standardised testing and teacher assessments. However, some schools are moving towards alternative forms of assessment such as the use of open-ended tasks and the development of student profiles. In most cases a variety of assessments were being used.

With the use of SOS a slight trend was noted away from the use of teacher assessments and standardised testing and toward the use a unified system of assessment and profiles for the purpose of accountability. As noted in one of the

interviews, standardised testing had been maintained for the purpose of the justification of the school's results rather than as an assessment in its own right.

It was shown in the survey that city schools were more likely to use teacher assessments and less likely to develop unified systems of assessment than the country schools to demonstrate their accountability. Many schools both in the city and country areas used standardised testing, but few from either area were likely to use a profile based on the outcome statements as part of their MIS (Management Information System).

To monitor the development of students, teacher tests and standardised tests are again the most frequently used, although the use of standardised tests is slightly lower for the pilot schools. Anecdotal records and checklists are used by a considerable number of schools both pilot and non-pilot, in schools using SOS and those which are not.

Portfolios are used by a greater percentage of pilot than non-pilot schools as are student journals and self-assessments. Portfolios are used more by the non-pilot schools that are using SOS than by those that are not, although there is little difference in their uses of journals and student self-assessments.

Principal's tests have little acceptance.

Pilot schools on average use slightly more assessments (5.6 per school) to monitor the progress of their students than do the non-pilot schools (4.1), however, most schools would agree that a variety of assessment types is preferable.

The need for open-ended tasks to be developed alongside the Maths SOSs was strongly advised by one school in the survey and another in the interview. The value of performance tasks is being realised in some schools.

As a result of SOS and accountability measures some schools (less than half of the non-pilot schools and just over half of the pilot schools) have introduced additional assessments. The most common introductions were standardised tests and portfolios in the non-pilot schools and student journals in the pilot schools.

Most forms of assessment were generally considered to be valid and reliable although it was commented that the purpose behind the assessment was important and that a variety of assessments helped to build a better profile of the student. Some concern was voiced about Principal's tests (which were shown to be seldomly used). A slight concern was shown about the use of standardised tests these concerns mainly coming from schools that are using SOS. A similar concern was shown about student self-assessments mainly by schools not using SOS.

A very high level of staff involvement was shown in the development of school responses to accountability and over half the schools acknowledged the involvement of parent groups. In some cases this was through the School-Based Decision Making Group (SBDMG).

The schools using SOS in mathematics, both pilot and non-pilot, tended to agree more that the *Improving and Reporting Schools' Performance* (EDWA, 1996a) document had assisted in staff communication and school/community communication (to a slightly lesser extent) than did the non-pilot non-SOS group of schools.

Efforts are being made by all but a few schools to inform the wider community of the developments in the assessment of mathematics mainly through the student's reports and school news letters with some P&C and SBDMG involvement.

The schools' responses to school accountability in the learning area of mathematics as set out in *Improving and Reporting Schools' Performance* (EDWA, 1996a) does not seem to be linked to the school's size or location. There does,

however, appear to be a link between the approach to accountability and the assessment techniques being used with the school's level of acceptance of the Student Outcome Statements in mathematics.

Conclusion

The establishment of a curriculum framework based on the Student Outcome Statements is mandatory for all Western Australian government schools by the year 2003, even though concerns regarding Outcome Based Education have been voiced (Ellerton & Clements, 1994). How effectively schools will be able to make this transition to Outcome Based Education within the next four or five years is uncertain. The schools that have already begun to make changes in this direction will probably be more able to meet this target than the many schools that have not as yet made any plans. The advice that was given in this study by schools that have begun implementation is to take it slowly so as to develop a full understanding of the changes and their implications not only in the ways that children are expected to develop and progress according to the SOS, but also in the teaching methods that will best facilitate the children's development. This will necessitate considerable changes in the planning of teaching programmes and methods of assessment, for most teachers, if the full potential of the planned changes is to be realised. Winnett and Gear (1997) on the basis of their own experiences with SOS suggest that a complete change of mind set will be needed and that teachers will be forced to confront their beliefs about how children learn, about their role as a teacher, and about how the aims of the programme and how the student's progress should be assessed. Large scale professional

development programmes and inservicing of teachers will be needed if the changes are to produce the desired effect in the classroom.

How these proposed changes will affect school accountability procedures will largely depend on the extent to which the individual schools adopt the changes. In turn this will depend on the philosophies of the teachers within the schools and how well the planned changes complement their own theories of how children best learn and how subjects should be taught. These philosophies and beliefs will cloud the individual's interpretation of the proposed changes. As McLeod states "beliefs and attitudes are relatively stable and resistant to change" (1990, p. 16), therefore we can not expect the implementation of SOS to have an immediate effect on the situation within individual classrooms.

The beliefs of the District Superintendents will also influence the schools decisions with regard to accountability as they impart their expectations to their schools. Accountability measures could vary considerably from District to District.

In the learning areas I feel that it would have been beneficial to have allowed the teaching staff of schools to become conversant with the SOS before the pressures of accountability were imposed. Now that schools have established their methods of assessing their accountability, these methods could also be resistant to change regardless of the changes implied by Outcome Based Education.

Whether SOS will be used to report to the parents of the students is a debateable issue, the main concern being that parents could become upset at their children's apparent lack of progress from one level to another even though there could be progress within the level. This system of reporting is not, however, new to the parents of children in Western Australia who have seen their children progress through

stages in their ability to swim, with the children assessed against specific criteria that are stated for the attainment of each swimming stage.

Many of the forms of assessment that are more compatible with the use of SOS ask the students to "take risks". In many ways the implementation of SOS and the accompanying accountability measures are asking the teachers to "take risks" too. These risks involve possible changes to the presentation of the curriculum, possible changes to the modes of assessment used, possible changes to the methods of reporting to parents, possible changes in the extent that teachers work and collaborate with their peers as they develop a shared understanding of what the outcomes actually mean and how the work samples collected from the students reflect these outcomes. How teachers and schools respond to these challenges is still to be seen as schools begin to plan and to implement their plans. The results of this study have shown that there are some changes being made in schools where Student Outcome Statements are being used and this in itself is promising for the future of Student Outcome Statements in this state.

References

Australian Education Curriculum Council (1993a). *A national statement on mathematics for Australian schools*. Carlton Vic: Curriculum Corporation.

Australian Education Curriculum Council (1993b).*Mathematics - A curriculum profile for Australian schools*. Carlton Vic: Curriculum Corporation.

Brandt, R. (1992). A fresh focus for curriculum, *Educational Leadership*, 49(8), 2.

Brandt, R. (1992). On performance assessment: A conversation with Grant Wiggins, *Educational Leadership*, 49(8), 35-37.

Burns, R. B. (1990). *Introduction to research methods*. Melbourne: Longman.

Clarke, D. (1988). *Assessment alternatives in mathematics*. Canberra: Curriculum Development Corporation.

Clarke, D. J., Clarke, D. M., & Lovitt, C. J. (1990). Changes in mathematics teaching call for assessment alternatives. In T. J. Cooney & C. R. Hirsch (Eds.), *Teaching and learning mathematics in the 1990s*. Reston, VA: NCTM.

Clarke, D. (1997). Anecdotal records interviews. In A. J. Herrington, L. Sparrow, J. Herrington & R. Oliver, *Investigating assessment strategies in mathematics classrooms*. CD-Rom resource. Perth, WA: MASTEC.

Clements, M. A. & Ellerton, N. F. (1995). *Research into the effectiveness of pencil-and-paper instruments for measuring mathematical understanding*. Paper presented to the 9th annual conference of the Educational Research Association. Singapore.

Chapman, L. R. (1972). Attempts to investigate the learning of mathematics. In L. R. Chapman (Ed.). *The process of learning mathematics*. Oxford, UK: Pergamon.

Cohen, L. & Manion, L. (1980). *Research methods in education*. Kent: Croom Helm.

Collis, K. (1989). Evaluation. In P. Langford (Ed.), *Educational psychology: An Australian perspective*. Melbourne: Longman Cheshire. 183-211.

Education Department of Western Australia (1989). *Learning Mathematics: Pre-primary to stage 7 mathematics syllabus*. Perth: Curriculum Programmes Branch, EDWA.

Education Department of Western Australia (1994). *Mathematics student outcome statements with pointers and work samples (working edition)*. Perth: EDWA.

Education Department of Western Australia (1995). *Making the links - A first steps in mathematics - Interim document*. Perth: EDWA.

Education Department of Western Australia (1996a). *Improving and reporting schools' performance (draft).* Perth: EDWA.

Education Department of Western Australia (1996b). *Mathematics report: Report of the student outcome statements trial, 1994 - 1995.* Perth: EDWA.

Education Department of Western Australia (1997a). *Western Australian Schools - Alphabetical List (Semester 1,1997)* Perth: EDWA.

Education Department of Western Australia (1997b). *School reporting: A framework for improving and reporting.* Perth: EDWA.

Ellerton, N. & Clements, M. A. (1994). *The National Curriculum Debacle.* Perth: Meridian.

Engelhart, M. D. (1972). *Methods of educational research.* Chicago: Rand McNally.

Garofalo, J. & Lester, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for research in mathematical education,* 16(3), 163-176.

Guskey, T. R. (1994). What you assess may not be what you get. *Educational Leadership,* 51(6), 51-54.

Howson, G. (1991). *National curricula in mathematics.* London: The Mathematical Association.

Hebert, E. A. (1992). Portfolios invite reflection - from students and staff. *Educational Leadership,* 49(8), 58-61.

Herman, J. L. (1992). What research tells us about good assessment. *Educational Leadership,* 49(8), 74-77.

Hillway, T. (1969). *Handbook of educational research: A guide to methods and materials.* Boston: Houghton Mifflin.

Johnson, M. C. (1977). *A review of research methods in education.* Chicago: Rand McNally.

Kissane, B. (1997). Outcomes and monitoring standards in education. In J. Gear, J. Johnston & B. Kissane (Eds.), *Using student outcome statements in mathematics: the NPDP experience.* Claremont, WA: MAWA.

Knight, P. (1992). How I Use Portfolios in Mathematics. *Educational Leadership,* 49(8), 71-72.

Larter, S. & Donnelly, J. (1993). Toronto's Benchmark Program. *Educational Leadership,* 50(5), 59-62.

Leder, G. (1991). Assessment: As we sow we reap. In M. Stephens & J. Izard (Eds.), *Reshaping assessment practices: Assessment in the mathematical sciences under challenge*. Hawthorn, Vic.: ACER.

Lofty, J. S. (1993). Can Britain's national curriculum show America the way? *Educational Leadership*, 50(5), 52-55.

Madaus, G. F. & Kellaghan, T. (1993). The British experience with authentic assessment. *Phi Delta Kappan*, 74(6), 458-469.

Marzano, R. J. (1994). Lessons from the field about outcome-based performance assessments. *Educational Leadership*, 51(6), 44-50.

McLeod, D. B. (1990). Information-processing theories and mathematics learning: The role of affect. *The International Journal of Educational Research*, 14(1), 13-29.

Ministry of Education. (1993). *The common curriculum grades 1-9*. Ontario, Canada: Ministry of Education and Training.

Nicholls, J. G., Cobb, P., Yackel, E., Wood, T. & Wheatley, G. (1990). Students' theories about mathematics and their mathematical knowledge: Multiple dimensions of assessment. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics*. Washington, DC: American Association for the Advancement of Science. 137-154.

Nisbet, J. D. & Entwistle, N. J. (1970). *Educational research methods*. London: Unibooks.

Nuttall, D. L. (1992). Performance assessment: The message from England. *Educational Leadership*, 49(8), 54-57.

O'Neil, J. (1992). Putting performance assessment to the test. *Educational Leadership*, 49(8), 14-19.

O'Neil, J. (1993). On the new standards project: An interview with Lauren Resnick and Warren Simmons. *Educational Leadership*, 50(5), 17-21.

Redding, N. (1992). Assessing the big Outcomes. *Educational Leadership*, 49(8), 49-53.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, Florida: Academic Press.

Simmons, W. & Resnick, L. (1993). Assessment as the catalyst of school reform. *Educational Leadership*, 50(5), 11-15.

Spady, W. (1993). *Outcome-based education*, Workshop Report, 5. Canberra: Australian Curriculum Studies Association.

Spady, W. G. (1994) Choosing outcomes of significance. *Educational Leadership*, 51(6), 18-22.

Szetela, W. & Nicol, C. (1992). Evaluating Problem Solving in Mathematics. *Educational Leadership*, 49(8), 42-45.

Telstra (1997). *White pages (Perth).*

Wiggins, G. (1992). Creating tests worth taking. *Educational Leadership*, 49(8), 26-33.

Willis, S., & Kissane, B. (1995a). *Outcome-based education: A review of the literature.* Report for the Education Department of Western Australia.

Willis, S., & Kissane, B. (1995b). *Systemic approaches to articulating and monitoring expected student outcomes*. Report for the Education Department of Western Australia.

Winnett, S. & Gear, J. (1997). Change processes in a primary school. In J. Gear, J. Johnston & B. Kissane (Eds). *Using student outcome statements in mathematics - The NPDP experience.* Claremont, WA: MAWA.

Wolf, D. P., LeMahieu, P. G. & Eresh, J. (1992). Good measure: Assessment as a tool for educational reform. *Educational Leadership*, 49(8), 8-13.

Yackel, E., Cobb, P. & Wood, T. (1992) Instructional development and assessment from a socioconstructivist perspective. In G. Leder (Ed). *Assessment and learning of mathematics.* Hawthorn, Vic: ACER.

A Survey To Determine How Schools Are Responding To The Requirements Of Accountability In Mathematics Education.

Your response to this questionnaire is greatly appreciated.

1. The number of full time teaching staff.....
   The number of part-time teaching staff.....

Please tick the most appropriate response box for each of the following question. If you would like to clarify any of your responses or add any personal thoughts about the question please use the section marked 'further comment'.

2. Which Learning Areas Student Outcome Statements are being developed and are being used throughout your school?

The Arts	❑	Health and Physical Education	❑
English	❑	Studies of Society and Environment	❑
Mathematics	❑	Languages other than English	❑
Science	❑	Technology and Enterprise	❑

Further comment:.....
.....
.....
.....
.....

3. Which additional Learning Area Student Outcome Statements are being used by at least one of the teachers in your school?

The Arts	❑	Health and Physical Education	❑
English	❑	Studies of Society and Environment	❑
Mathematics	❑	Languages other than English	❑
Science	❑	Technology and Enterprise	❑

Further comment:.....
.....
.....
.....
.....

4. When does your school plan to adopt the use of the Student Outcome Statements in Mathematics throughout the school?

Already in use ❑ 1998 ❑ 1999 ❑

2000 ❑ Has not been decided ❑

Any comments on the introduction of Student Outcome Statements:
.....
.....
.....
.....
.....

5(a). Is there a copy of the document *Improving and Reporting Schools' Performance* (EDWA, 1996) in the school?

Yes ❑ No ❑

(b). Is there a copy of the document School Performance: A Framework for Improving and Reporting (EDWA, 1997) in the school?

Yes ❑ No ❑

(c). Have the school staff discussed the implementation of these documents?

Yes ❑ No ❑

Further comment:.....
.....
.....
.....
.....

6. Is there a 'specialist' mathematics teacher in your school?

Yes ❑ No ❑

Further comment:.....
.....
.....
.....
.....

7. Which statement(s) best describe(s) your school's approach to accountability in mathematics?

Each teacher assesses their own mathematics programme in their own way. ❑

A unified system of assessment has been adopted throughout the school to show each child's progress. ❑

Standardised testing is conducted throughout the school to monitor progress. ❑

A profile, based on the outcome statements, has been established to monitor each child's progress throughout their primary schooling. ❑

Others (please specify):.....
.....

Further comment:.....
.....
.....
.....
.....

8. Which forms of assessment are used to monitor the students' development in mathematics education?

Teacher tests	❑	Student journals	❑
Checklists	❑	Standardised tests	❑
Portfolios of work samples	❑	Student self-assessments	❑
Anecdotal records	❑	Principal's tests	❑

Others (please specify):.....
.....

Further comment:.....
.....
.....
.....
.....

9. Which of these forms of assessment have been introduced as a result of either the Student Outcome Statements or the Improving and Reporting Schools' Performance documents in mathematics education?

Teacher tests	❑	Student journals	❑
Checklists	❑	Standardised tests	❑
Portfolios of work samples	❑	Student self-assessments	❑
Anecdotal records	❑	Principal's tests	❑

Others (please specify):.....
.....

Further comment:.....
.....
.....
.....
.....

10. Which of these forms of assessment do you think do not give valid and reliable information regarding student progress in mathematics education?

Teacher tests	❑	Student journals	❑
Checklists	❑	Standardised tests	❑
Portfolios of work samples	❑	Student self-assessments	❑
Anecdotal records	❑	Principal's tests	❑

Others (please specify):.....
.....

Further comment:.....
.....
.....
.....
.....

11. Which of the following parties have had input into how the school has approached accountability?

The Principal	❑	The deputy principals	❑
Some teaching staff	❑	All teaching staff	❑
Parent groups	❑		

Others (please specify):.....

Further comment:.....
.....
.....
.....

12. The *Improving and Reporting Schools' Performance* document has caused an increase in communication between the members of the teaching staff in regard to the assessment of mathematics.

Strongly Agree	❑	Agree	❑
Undecided	❑	Disagree	❑
Strongly Disagree	❑		

Further comment:.....
.....
.....

13. The *Improving and Reporting Schools' Performance* document has caused an increase in communication between the members of the teaching staff and the wider school community in regard to the assessment of mathematics.

Strongly Agree	❑	Agree	❑
Undecided	❑	Disagree	❑
Strongly Disagree	❑		

Further comment:.....
.....
.....
.....