

Planning valid and valuable school assessment systems: A possible model for New Zealand schools

Mary Hill

Department of Professional Studies, University of Waikato

Introduction

My recent work in the field of assessment has brought me into close contact with many schools wrestling with assessment policies and practices. This will come as no surprise to most of you. The production and implementation of the national curriculum, and the promulgation of regulations governing how schools must implement it and monitor students' progress through it (e.g. the National Education Guidelines - NEGs' and the National Administration Guidelines - NAGs), has continued remorselessly over the last three years. Schools and teachers have borne the brunt of working out in practice (i) how assessment will take place, (ii) what assessment activities are the most valid and useful, (iii) how the results can be used for informing students and their parents about their progress and (iv) how, as a school, they will demonstrate to the community at large, and the Education Review Office (ERO) in particular, that the school is making a difference to children's learning.

Following devolution of responsibility for management to the school level, each school has had to work through, with few guidelines and often little support, how to operationalise assessment in the new curriculum environment. Of course, many of the assessment practices existing before tomorrow's schools and the introduction of the new national curriculum documents have survived into the new environment. The child-centred and individualised feedback about the quality and amount of learning promoted in *Keeping School Records* (Department of Education, 1989) has become so much part of the New Zealand primary school teacher's way of functioning that it is now a taken-for-granted part of the teaching process. Standardised tests, such as the Progressive Achievement Tests (PATs) and Running Records in reading, also remain as tools in our assessment kits. Less familiar to New Zealand schools and teachers, however, are ways to provide information to judge the effectiveness of schools as institutions. Due to the expectations of the Education Review Office for schools to produce this information, and the interest of the media in publishing it, pressure has mounted on principals and teachers to develop appropriate ways to undertake this type of assessment.

In 1994, Dr. Ian Calder and I carried out a survey of the assessment knowledge and confidence of over two hundred teachers in the Waikato and the Bay of Plenty area.

As well, interviews carried out by our assessment course students illuminated the survey results and confirmed that

- teachers lacked confidence and knowledge in the assessment area,
- felt they were being asked to meet increased assessment demands, and
- believed that the new assessment environment had increased their workload, accountability demands and had led to a greater emphasis on record keeping.

Since 1994, I have been involved in facilitating teacher development in assessment at the school level. I have found that two factors lead to a more confident approach. These are (i) helping teachers understand a few central assessment concepts, and (ii) promoting communication about assessment practices between staff, school management and the school community. The purpose of this paper is to outline the assessment ideas teachers have found useful and describe how two schools have used the model provided to implement valid and valuable assessment systems.

The Roles of Assessment in Schools

Within schools, assessment serves several clearly defined purposes. I have found that teachers frequently dismiss, as jargon, the classification of assessment into these purposes. This is unfortunate because a strong understanding of the purposes of assessment can greatly assist teachers to construct valid and valuable systems for supporting learning, reporting to parents and assuring themselves and others about the quality of the education being provided within their school. My suggestion is that rather than jargon, teachers come to think of the following terms as part of their professional language through which they can accurately communicate with other teachers and educators, and which they can refer to when building a model that describes how various assessment activities complete the assessment picture in their school.

Firstly, assessment provides feedback to teachers and pupils about on-going progress in learning. It has a direct influence on the quality of pupils' learning experiences and thus on the level of attainment that can be achieved. This type of assessment is known as **formative** assessment. In contrast, assessment that is the means for communicating the nature and level of pupils' achievements at various points in their schooling is called **summative** assessment. A third role, known as assessment for certification, is used to summarise for the purposes of selection and qualification. Lastly, assessment that provides part of the information used in judging the effectiveness of educational institutions and of the system as a whole is called **evaluative** assessment or **quality control**. (Harlen, Gipps, Broadfoot and Nuttall, 1992).

In primary schools and, in particular, junior school departments, only three of these roles for assessment are generally used. Formative assessment is far and away

the most common. Summative assessment is also prominent and often confused with the quality control functions of assessment. Assessment for certification is rarely if ever used in primary schools so is not referred to again in this article.

Figure 1 below provides a theoretical framework upon which schools can build, and indeed have built, a practical system of assessment. Examples of how this can be done are suggested in the next section of this article. It is possible, of course, to use a very large variety of strategies to fulfil the purposes of assessment referred to in this model. The examples offered in Figures 2 and 3 are just that, two possible ways of designing a school-based assessment system. The model in Figure 1 is offered as a guiding framework that teachers can use to explain how each purpose of assessment is met within their own school's system and how each part of their system is related to the whole.

Formative Assessment

The assessment described to me most often by teachers as of greatest value is formative assessment. For example, when asked what assessment she valued most, one teacher in my current research project stated that it is:

...the assessment activities where I'm working on a one-to-one [basis] with children. During that time it's not only what the child actually does in response, but I also find it's how the child reacts to it, because from there I can glean more information from the non-verbal as well as the verbal responses - positive reactions, negative reactions. It's even as simple as how the child holds his/her pencil, whether his or her coordination is hindering the process.

Another said it was:

providing you with a means of taking a child on to a next stage; eyeing accurate data on where that child is in relation to your objective.

In these descriptions, as in others I have collected, the teachers are describing formative assessment. Such assessment is an integral part of the interactions between teacher, pupil and learning materials. Because of this relationship, some teachers who use formative assessment well may not even recognise that what they are doing includes assessment. Probably this is because they think of assessment as a more formal activity, distinct from teaching (Harlen et al, 1992).

During, or as a result of, formative assessment, most teachers record in some way what they are finding out. They do this in order to remember what they have observed as an ongoing record of progress, and as an aid to further teaching. The impression I am gaining from my interactions with teachers is that they only record significant and useful information. Because recording is a time consuming process, they are selective and usually record information focussed on aspects that indicate

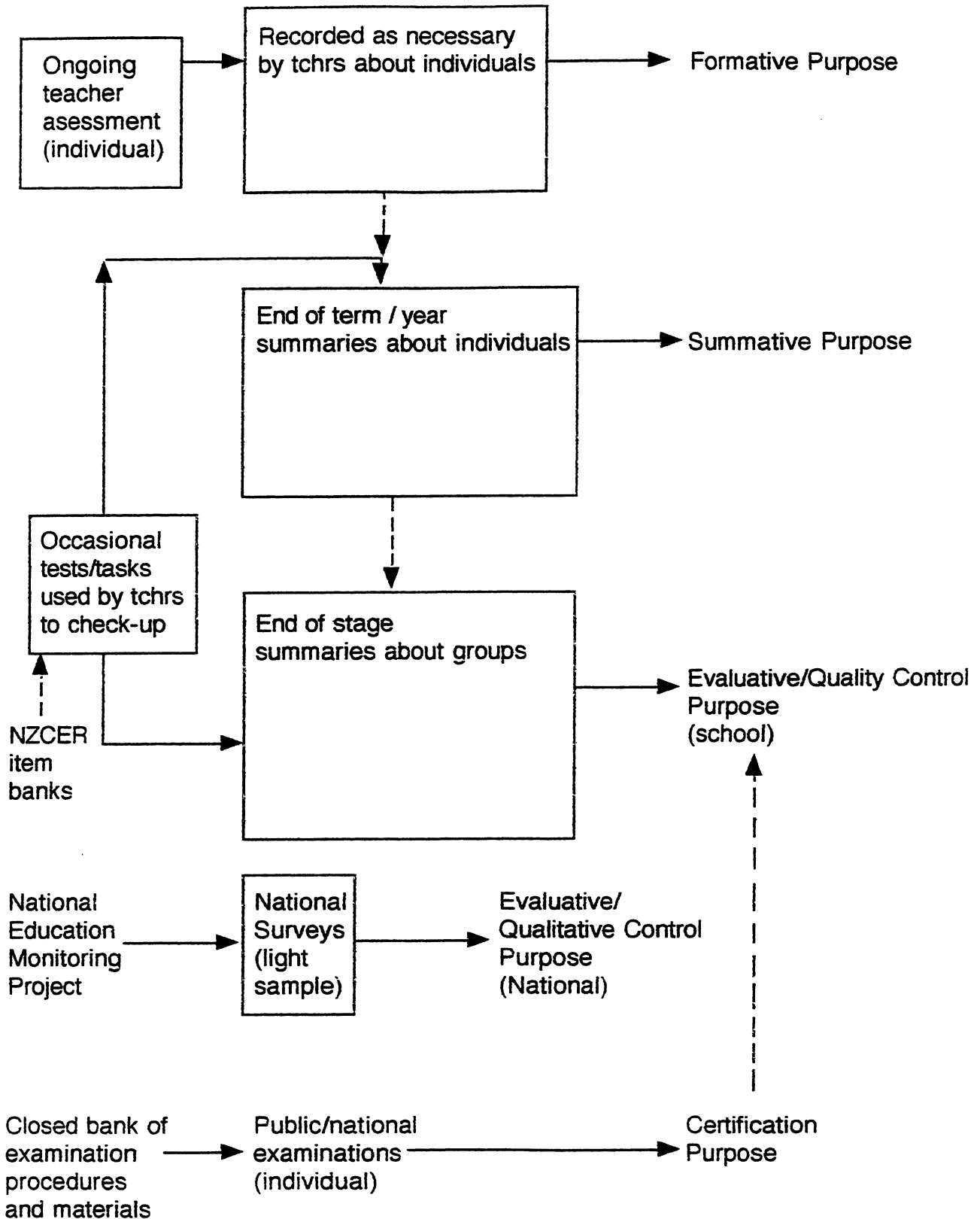


Figure 1. Theoretical Assessment Model (Adapted from Harlen et. al., 1992)

certain needs or strengths. These recordings are, importantly, the **clinical case notes** of the professional teacher. Further, my investigation is revealing that, because these **field notes** are so crucial for focussing teaching, as well as for planning and grouping, teachers tenaciously stick to their own personally developed methods for this initial type of recording.

Some teachers are very well organised in the collection of these 'case notes' and set up well-organised 'roll books' with the major learning objectives along one axis of the pages and the children's names down the other. Cutting the page edges to reveal the different areas of the curriculum, as in an address or telephone directory, allows quick access to them. A junior school teacher explained her use of such a document to me in this way:

So it's with me , nice and close to me whenever I want it. Participation and everything. And notes, and I find now, that because I know the document [her roll book] so well, ... that as the children are doing their work, I can quickly just note down in it what's happening or what they need. I have it on hand. I couldn't work without it. I need to refer back every so often.

Other teachers use other recording mechanisms, including exercise books, clipboards, clear files and ring binders to manage their note taking. But all lend themselves to helping the teacher gather the information he or she requires regarding progress towards the main goals of the teaching programme.

The information that is gathered during teaching is closely related to the purposes for teaching. It is information that tells about children's progress toward the criteria the teacher has in mind while he/she is teaching, and it is recorded about each student, firstly for the purpose of deciding what further learning is required and secondly, as an ongoing record of progress. Generally, this information is not aggregated to show the achievement of groups of students but it is often shared by the teacher and the student and can be used for reporting to parents.

Because the information is collected during actual teaching and learning, it usually has high validity. In simple terms this means that the information gathered relates very closely to what you planned to find out (Sutton 1992). Teachers then usually use this information to assist students' learning, another important aspect of validity (Crooks, 1996, Sutton, 1992). Such high levels of validity cannot always be credited to other kinds of assessment, including summative assessment.

Summative Assessment

Summative assessment, like formative, also concerns the performance of individuals as opposed to groups of students (Harlen et al, 1992). Its prime purpose, however, is not to influence teaching (as in the case of formative assessment), but rather to summarise information about the achievements of a student at a particular time.

Harlen et al (1992) neatly classify the twin reasons for gathering summative information as “summing up” and “checking up” (p.222). Summing up is some form of summary of information obtained through recording formative assessments during a particular period of time, while checking up is the collection of new information about what a student can do at the end of a period of time, often using an assessment task or perhaps a test.

This is the point at which confusion often occurs for the teachers I work with. Rather than summarising the information that has been collected during formative assessment to give a broadly based but brief portrait of the student’s achievement, many schools have developed detailed objective by objective profiles about individual students which duplicate the information captured in the field/case notes described above. When asked why this is done, the main reason given is a genuine attempt to meet the needs of people other than the teacher and the student. For example, these profiles are thought to be for the senior teachers and the principal, for the parents, for teachers at other schools or levels, and for the Education Review Office. In practice, however, teachers are reporting that these other people are not using the information. The comment one teacher made to me was common:

I question who’s going to pick them up and go through and see a particular child has achieved in science - whether they know a certain objective, I mean who is going to look at it and what use is it going to be? I mean ... I get the profiles at the beginning of the year and the things I look at ... are their reading ages to find out where I can start them, and their maths levels ... and I look at their social comment. After a few weeks ... having an opinion of my own ... if I’ve got a concern, I’ll compare it with the teacher last year. But apart from that, I don’t particularly look at them. I think it’s being done because we’re here to please outside agencies and we need to be seen to be doing it to show our worth I guess.

This teacher voices a commonly held view that the summative information is mostly for quality control or evaluative purposes, rather than being a way to summarise information about particular students at a particular time.

In terms of the “checking up” aspect, primary teachers use regular teacher-made tasks and standardised procedures, particularly in the areas of literacy and mathematics. Most schools use nationally standardised tests at school entry (Ministry of Education, 1992), at six years of age (including a diagnostic survey of reading and written language, Clay, 1993) and throughout the rest of primary schooling (PATs are commonly used). In addition, many schools have policies for the use of further tasks, checkpoints and the collection of work samples that assist teachers to keep an eye on expected student achievement. The results of using these instruments are also commonly recorded and collected in the individual student profile, frequently making this document large and weighty. Teachers have described carrying large cartons of

records backwards and forwards between school and home in order to record the required data.

On a positive note, many schools are currently reviewing their assessment systems and rationalising the record-overload that has resulted from a lack of clarity of why summative assessment is necessary and who it is for. The main purpose of summative records can be achieved with a **cumulative record** of the “checking up” results, and a summary of the progress recorded through formative assessment. These records stand as the ‘point in time’ documents for each pupil so that new teachers can pitch programmes appropriately, senior teachers and principals can gain a ‘bird’s eye view’ of progress and teachers, students and parents can see progress graphically. In the model represented in Figures 1, 2 and 3, the line from the formative recording box to the summative recording box is dotted to indicate that only a summary of the information gathered from formative assessment is recorded for summative purposes. The solid line from ‘occasional tests/tasks used by teachers to check up’ to both the ‘end-of-term/year summaries about individuals’ and the ‘end-of-stage summaries about groups’, indicates that this information can be used for both purposes. Teachers within each school will need to make decisions about how to use the information gained from these sources based on their purposes for using them and their knowledge about how such results can be used.

Assessment for Evaluative and Quality Control Purposes

As a result of self management there has been an increased emphasis on schools collecting information to show that students are attaining desired levels of achievement. The surveillance of schools by the Education Review Office and the possibility that the results of such reviews will be published in the media, has resulted in evaluative assessment becoming a high-stakes activity. A great deal of anxiety about this type of assessment exists among teachers. One teacher told me:

When I say assessment, it's like checking ... things that are there to please other people and not me. The big ERO thing ... over your head really. You've got someone who tells you ... like your team leader you have to have this because someone's told them because higher up ... eventually it gets to the ERO system peering over us to check that we're doing our job. And you can never do enough to please ERO. This is the sort of feeling that's out there ... ERO are going to come in here, they are going to look at your school, they are going to assess you, it's going to be in the paper, if it's bad you don't get kids coming to your school and you're left, you don't have a job.

The problem here is that too often teachers believe that they should use the information collected for formative purposes to meet school quality control requirements. This sets up an inevitable tension between high-stakes assessment for accountability and low-stakes assessment for better learning; the result is schools

struggling to get things right (Black, 1995, 1996). As alluded to in the above quotation, the role of teachers at different levels of the school influences the importance attached to each kind of assessment. At the classroom level, the teacher needs assessment for better instruction, while at the management levels of principals, assistant and deputy principals, evaluative assessment and accountability are crucial to enable targeting of resources and reporting on the school's performance (Elley, 1996). Sorting out how these needs will be met through the efficient and effective use of assessment is the challenge that confronts most schools and teachers at present.

The information used in assessing the effectiveness of schools should communicate a wide range of information featuring the achievement of students that is attributable to the effect of the school rather than other factors (ERO, 1995; Harlen et al, 1992). Harlen et al (1992) caution that a model of evaluation is required that disentangles the effect of attainment from that of the pupil's background. They support an approach which looks at the gain in achievement while the pupil is at a particular school, that is, the progress she/he makes there. In this approach, they recommend that,

assessment of attainment used at entry to the particular school and at various point after that, should be broad as possible to ensure that school effectiveness is not reduced to efficiency in teaching test-taking skills but reflects the full range of the aims of the school (p.227).

In order to do this, schools need to articulate what it is they want their students to achieve. This is not as straight forward as it may sound. ERO (1995), in a recent report about assessing student achievement, stated that while this approach has the potential to provide information on students' progress, most schools had not clarified achievement expectations, often did not have baseline information, did not yet have access to reliable measurement instruments, and are restricted in the number of students for whom achievement could be measured due to high student mobility. Harlen et al (1992) and Flockton (1996) are more positive about the possibilities. With a statement of expected achievement clarified, assessment information about the effectiveness of a school could be published in the form of reports including:

- the aims of the school
- details of recent reviews
- particular areas of expertise offered
- cultural and sporting achievements
- community involvement
- academic achievement in relation to the stated aims

Importantly, assessment for quality control and evaluation purposes should not drive the rest of the assessment system in the school. It should be mainly the responsibility of senior management within the school, take only a small proportion of assessment time in comparison with assessment for the other purposes stated above,

and could include sampling techniques if the school is big enough. In the end, the quality of the formative assessment carried out and the teachers' ability to improve teaching and learning as a result will be the most important factor in raising the achievement levels, not the assessment carried out for quality control purposes.

The model for school-based assessment (Figure 1) attempts to show graphically with a dotted line how some of the information gathered for summative purposes may be used for evaluative/quality control purposes (where teachers decide this is appropriate) along with information they collect from other sources. In the future, information from the NZCER item banks may contribute to both summative and evaluative purposes at Years 7 and 9. Although not indicated in the model presented in this paper, other sources of descriptive information, such as photographs of the results of a whole-school gardening project or the enthusiasm generated by a cultural concert, will most likely form part of the data meeting the evaluative/quality control function of assessment.

At the national level in New Zealand rather than standardised or nation testing, the evaluative/quality control purpose of assessment if being undertaken through light sampling techniques carried out under contract to the Ministry of Education by the National Education Monitoring Project (Crooks and Flockton, 1995). This is represented on the model (see Figure 1) by a separate category underneath the school-based elements of the model.

Finally, external examinations, carried out in the last three years of high school are included in the model to show how the certification purposes of assessment are related to the rest of the system. The dotted line from this purpose to the 'school evaluative/quality control' purpose indicated that many secondary schools also use the results of these external exams as part of the data they use to report on their quality.

Two Examples of School-Based Assessment Systems

The theoretical framework described in this article underpinned the assessment systems designed by two primary schools with which I have worked recently. In the first example (see Figure 2) the teachers made a decision to use 'roll books' in which to record field notes about the information they gained through formative assessment. The teachers agreed that they could organise and complete these in any way they wanted, as long as they covered the national curriculum essential learning areas and skills relevant to their class levels. This did not mean checklists of the achievement objectives, but field notes about the performance of individuals in relation to the curriculum objectives that the teachers themselves had decided were important.

The teacher at this school resolved to reduce the information they kept in folders for each individual child so that they contained

- a four-sided tracking sheet to cover the first six years of schooling in all

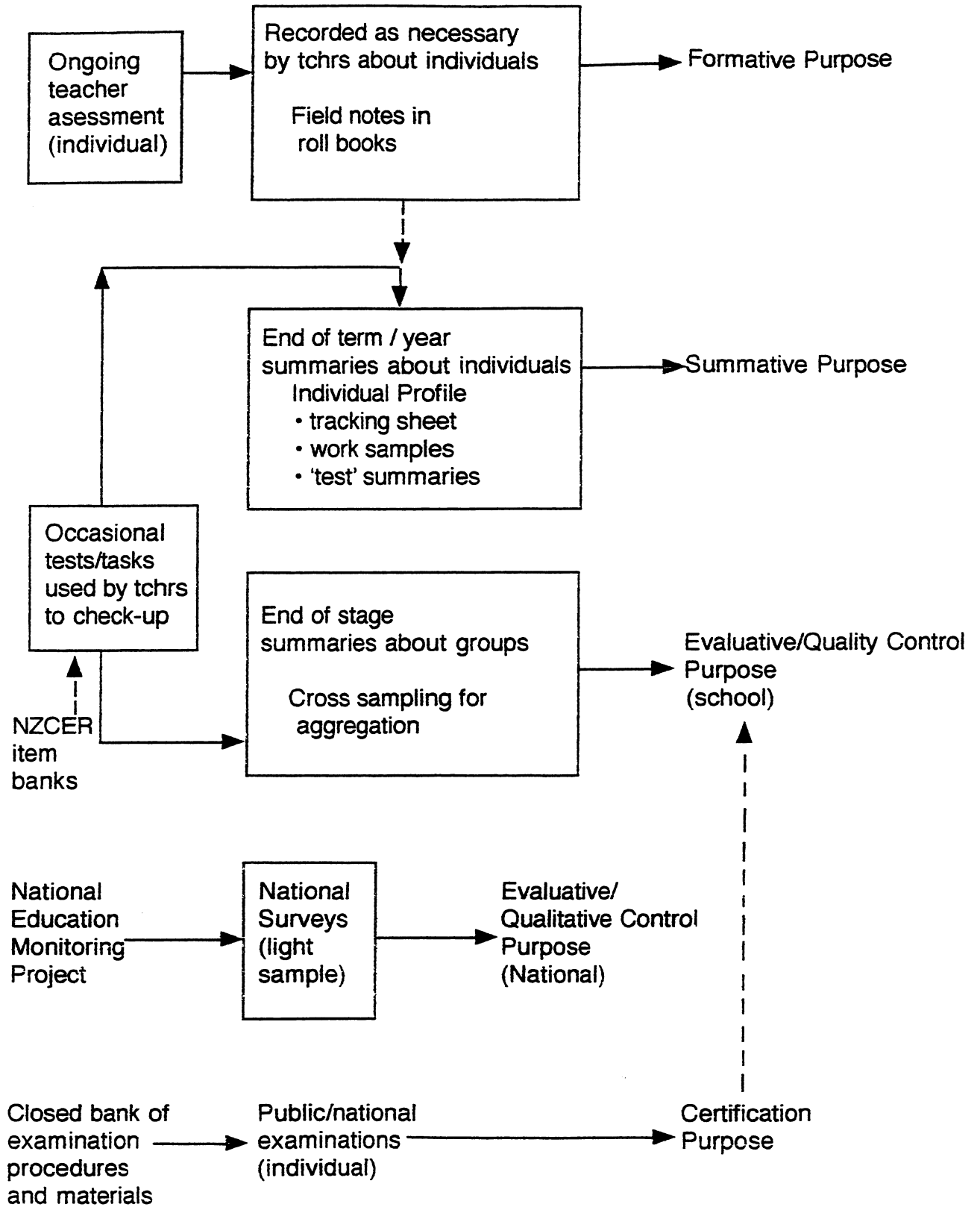


Figure 2. Theoretical Assessment Model : Example A

curriculum areas,

- appropriate work samples, particularly in written language, and
- any test summaries which they agreed were necessary, for example, the results of the six year diagnostic survey in literacy.

At this school, to monitor the quality of learning within the school, they had already implemented a system of sampling groups of children representative of different abilities and gender/ethnic groups. They decided to continue this practice to meet the evaluative/quality control function of assessment.

The second example (see Figure 3) a school-based assessment system allowed the teachers much more freedom to collect information from their formative assessments. Teachers could use whatever system they desired to record data about all essential skills and learning areas, and were provided with two differently formatted books to use if they wished. However, they had decided to trial a student portfolio of work samples in 1996 in which the children were required to store particular samples of their work. These samples were, in some cases, used by the teacher for formative assessment purposes. The work samples in the portfolios were also used as a way of monitoring progress throughout the year and as a vehicle for communicating with parents about their child's progress and achievement.

The school in example B also designed a four-sided tracking sheet onto which progress in all curriculum areas was summarised throughout primary schooling, including the results of standardised tests. Some of the data from the formative assessment process and from the summative records was used to meet the evaluative/quality control purposes of assessment and to indicate possible barriers to learning, but other information from additional assessment tasks, schoolwide initiatives (such as the gardening project mentioned above), and the students' own portfolios was also used to demonstrate how well the school had achieved its stated purposes.

Some Unresolved Issues

The new national curriculum has been arriving in schools now since 1993. Consequently, there has been time for some curriculum inservice work and even some limited teacher development in assessment practices. Understanding the different purposes for assessment can go some way towards assisting teachers reduce the anxiety attached to assessment and improve their practices. However, schools and teachers are still struggling with a range of issues in their attempts to clarify and improve assessment. Among these issues appear to be

- lack of teacher knowledge and confidence about assessment, and
- variation in the type of information gathered by teachers, depending on their views about learning and the uncertainty about the future expectations of outside agencies.

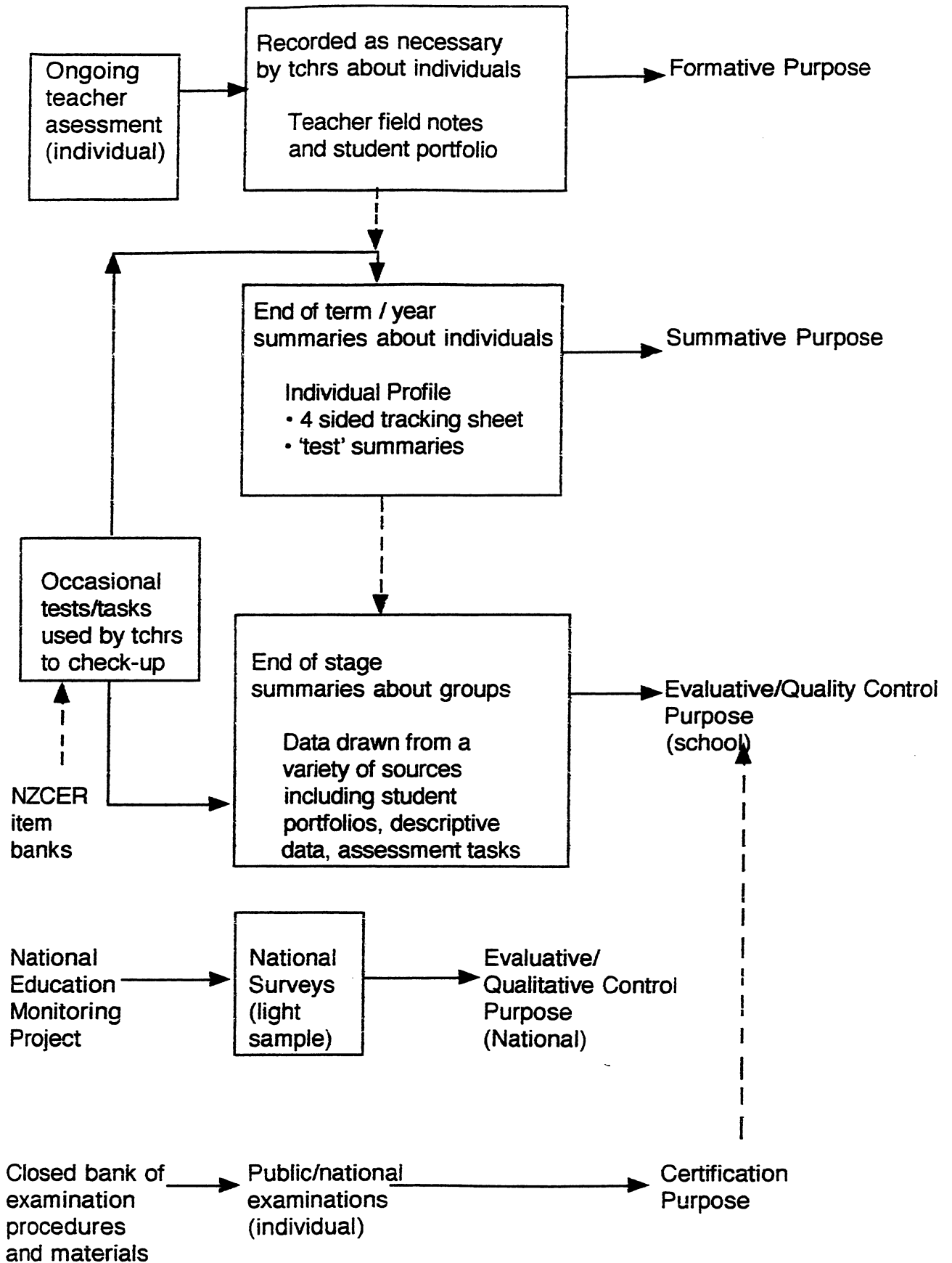


Figure 3. Theoretical Assessment Model : Example B

As mentioned above, a survey carried out in the University of Waikato region two years ago found that most teachers had had little formal training in assessment and most had learnt what they know 'on the job'. Nearly three-quarters of the teachers interviewed reported that their training had been poor. Of the 206 teachers surveyed, 147 described their understanding of formative assessment as low, 135 stated their understanding of summative assessment as poor, and 98 were not sure about validity or reliability. Despite a few teacher development programmes in assessment in the interim, my interactions with teachers would suggest the situation is still very much the same. There is an urgent need for more teacher education in assessment. At our University, assessment courses are now compulsory for primary preservice teachers. New courses in classroom assessment are available at both the undergraduate and graduate level but a very small number of practising teachers can take advantage of these. Wider-reaching strategies are needed to equip teachers with the knowledge and experience they need to work through the tensions and problems with assessment that presently exist in all schools (ERO, 1995, Harlen et al., 1992).

Associated with teachers' lack of knowledge, and therefore confidence, about assessment is the way teachers view learning. In some views of learning and teaching, assessment is the identification of what is and isn't known, so that the gaps might be filled and existing knowledge added to. However, when learning is considered to involve the teacher taking a child's initial ideas seriously to ensure that any change or development of these ideas, and the supporting evidence for them, makes sense and becomes owned by the child (Harlen, et al., 1992, p.220), much more than objective ticking is required. The way in which a curriculum document is structured can influence these views of teaching and learning. A document containing a strict hierarchy of small, supposedly sequenced objectives may influence teachers toward a checklist mentality. The proliferation of student profiles containing lists of objectives to be ticked off and dated on attainment is an indication that this has happened to some extent with the introduction of the new national curriculum.

Conclusion

To ensure valid and valuable assessment, schools and teachers within them need informed, confident teachers; teachers who know how to collect valid assessment information and use it appropriately. My hope in presenting this model and the accompanying examples, is that in clarifying the different purposes for which assessment is used and in identifying how teachers might go about achieving each of these assessment purposes, it will be easier to see the way forward. It is up to teachers in individual schools to work together to design and trial appropriate assessment practices, and to find out what works for them. Quite clearly, there is to be no national system of assessment and even teacher development in this area is mainly the

responsibility of individual schools. It is through school-based research and development of assessment practices and systems, such as the study from which quotations in this paper were drawn, that progress towards better assessment and improved teaching and learning will be made.

References

- Black, p. (1995). Can Teachers Use Assessment To Improve Learning? *British journal of Curriculum and Assessment*, 55, 2, 7-11.
- Black, P. (1996). Address to the International Symposium on Validity in Assessment held at the University of Otago, St Margaret's College, Dunedin, N.Z., 28-30 June.
- Clay, M. (1993). *An Observation Survey of Early Literacy Achievement*. Auckland: Heinemann.
- Crooks, T. (in press). Threats to the Valid Use of Assessments. To appear in *Assessment in Education: Principles, Policy and Practice*.
- Crooks, T. and Flockton, L. (1995). The National Education Monitoring Project: Developing Assessment Procedures and Tasks for New Zealand. *Best of SET: Assessment*, No.2.
- Department of Education (1989). *Keeping School Records: Primary Progress records*. Wellington: Government printer
- Elley, W. (1996). Standards Based Assessment and the New Zealand National Curriculum. Paper presented at the International Symposium on Validity in Assessment held at the University of Otago, St Margaret's College, Dunedin, N.Z., 28-30 June.
- Education Review Office (1995). *Assessing Student Achievement*. National Education Evaluation Report, No 3, Autumn.
- Faire, M.F. (1994). *Recording and Reporting: An Exploratory Study*. An exploratory study carried out under contract to the Ministry of Education. Hamilton: Department of Professional Studies, University of Waikato.
- Flockton, L. (1996). Address to the International Symposium on Validity in Assessment held at the University of Otago, St Margaret's College, Dunedin, N.Z., 28-30 June.
- Harlen, W., Gipps, C., Broadfoot, P. and Nuttall, D., (1992). Assessment and the Improvement of Education. *The Curriculum Journal*, 3, 3, 215-230.
- Ministry of Education (1992). *A Survey of School Entry practices: How schools gather information on new entrants*. Wellington: Learning Media.
- Sutton, R. (1992). *Assessment: A Framework for Teachers*. London: Routledge.