The Characteristics of Formative Assessment in Science Education

BEVERLEY BELL, BRONWEN COWIE
School of Education, University of Waikato, Private Bag 3105, Hamilton, New Zealand

ABSTRACT: Formative assessment, in this article, is defined as "the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning." The findings of a two-year research project in New Zealand indicate that formative assessment has the following characteristics: responsiveness, sources of evidence, a tacit process, using professional knowledge and experiences, an integral part of teaching and learning, formative assessment is done by both teachers and students, the purposes for formative assessment, the contextualized nature of the process, dilemmas, and student disclosure.   © 2001 John Wiley & Sons, Inc. *Sci Ed* **85:**536–553, 2001.

INTRODUCTION

Formative assessment is increasingly becoming a focus in policy documents on educational assessment and in the professional development of teachers. The term "formative assessment" is not new, but is now being used in more detailed and specific ways. As this happens, there is a call for further research and theorizing on formative assessment (Black & Wiliam, 1998). This article documents some research that explored the current practice of some science teachers to describe the formative assessment being done in some New Zealand science classrooms.

There have been three trends in education that have highlighted the need for teachers to do formative assessment: continuous summative assessment, multiple purposes for assessment and teaching, and assessment for conceptual development. Each of these will be briefly discussed in turn.

Continuous Summative Assessment

One of the trends in educational assessment that has put the spotlight on formative assessment is the development of more valid assessment procedures. In the 1970s and 1980s, there was much criticism of the validity of summative assessments used in educational assessment, and, in particular, of the limitations of the validity of external testing and examinations (Keeves & Alagumalai, 1998). This included criticisms of the validity of assessment tasks, such as multiple choice questions, and criticisms of norm-referenced assessments, such as those for national qualifications. There was also criticism of the impact of high-stakes, standardized testing on school learning (Black & Wiliam, 1998). The responses to these criticisms can be summarized as a need to:

Correspondence to: B. Bell; e-mail: b.bell@waikato.ac.nz

- Assess a wider range of science learning outcomes, such as performance of investigation skills (Johnson, 1989) and multiple forms of thinking (Gardner, 1985).
- Use a wider range of assessment tasks (other than multiple choice tests, questions requiring short answers and essay questions); for example, portfolios (Duschl & Gitomer, 1997; Gitomer & Duschl, 1995, 1998) and performance-based assessment (Erickson & Meyer, 1998).
- Integrate assessment with the curriculum and assess in more authentic contexts (Tamir, 1998).

As these recommendations could not be achieved through external examinations or standardized testing alone, assessment by teachers (also called internal assessment) was seen as a way forward. Hence, an early use of the term "formative assessment" was to distinguish between continuous summative assessment by teachers in the classroom and summative assessment by external examiners, such as those who develop standardized tests and those who set and mark examinations for national qualifications. This continuous summative assessment by teachers was initially called formative assessment as it did enable some information on learning to be given to students and teachers in the course of the school year, although it was relatively coarse feedback. It has been called "weak formative assessment" (Brown, 1996). The questions often raised during discussions on continuous summative assessment are: how many separate assessments have to be recorded for the aggregated mark or grade to be reliable and valid; how best to store the multiple assessment documentation; how to aggregate the marks or grades; the problems with reducing many assessment results into one grade; and whether all the achievement objectives in the science curriculum have to be assessed and how often.

Multiple Purposes for Assessment

Another trend, which has highlighted formative assessment, was the trend toward multiple purposes for assessment. This trend was brought into sharp focus in the 1990s, when politicians and others wanting to hold educationalists accountable, looked to assessment to provide the information required for the accountability process. This added to the existing demands for assessment information by people who operate outside the classroom (e.g., caregivers, principals, school governing bodies, local or national government officials, awarders of national qualifications, selection panels for tertiary education programs, and employers).

In New Zealand, and elsewhere internationally, this trend toward using educational assessments for accountability purposes in addition to the existing purposes has highlighted the multiple purposes for assessment. These multiple purposes can include auditing of schools, national monitoring, school leaver documentation, awarding of national qualifications, appraisal of teachers, curriculum evaluation, and the improvement of teaching and learning.

There are three cornerstones of the accountability process: a prescribed set of standards, an auditing and monitoring process to ascertain if the standards have been attained, and a way of raising standards if low standards have been indicated in the audits. The "raising of standards" is often seen by policymakers as being achievable by a number of methods, including school-based assessment (Ministry of Education, 1993).

There is now a recognized need to clarify these multiple purposes for assessment. In a recent example, in the United States, a working party (National Research Council, 1999) identified three purposes of assessment:

Assessment has multiple purposes. One purpose is to monitor educational progress or improvement. Educators, policymakers, parents and the public want to know how much students are learning compared to the standards of performance or to their peers. This purpose, often called *summative assessment* is becoming more significant as states and school districts invest more resources in educational reform.

A second purpose is to provide teachers and students with feedback. The teachers can use the feedback to revise their classroom practices, and the students can use the feedback to monitor their own learning. This purpose, often called *formative assessment,* is also receiving greater attention with the spread of new teaching methods.

A third purpose of assessment is to drive changes in practice and policy by holding people accountable for achieving the desired reforms. This purpose, called *accountability assessment*, is very much in the forefront as states and school districts design systems that attach strong incentives and sanctions to performance on state and local assessments. (National Research Council, 1999, pp.1−2)

In New Zealand, some of these multiple purposes for assessment are said to be addressed by using school-based assessment, the purpose of which can be described as, "improving learning, reporting progress, providing summative information, and improving programmes" (Ministry of Education, 1994, pp. 7–8). These multiple purposes of school-based assessment are seen as giving rise to three broad categories of assessment: diagnostic, summative, and formative assessment.

Formative assessment is described in the policy document as:

... an integral part of the teaching and learning process. It is used to provide the student with feedback to enhance learning and to help the teacher understand students' learning. It helps build a picture of a students' progress, and informs decisions about the next steps in teaching and learning. (Ministry of Education, 1994, p. 8)

While the foregoing description of formative assessment could include continuous summative assessment, the research documented in this article specifically explored formative assessment as classroom assessment to improve learning (and teaching) during the learning.

Teaching for and Assessment of Conceptual Development

The third trend that has highlighted formative assessment is the development of views of assessment to match the views of learning, which recognize that each learner has to construct an understanding for her- or himself, using both incoming stimuli and existing knowledge, and not merely absorbing transmitted knowledge (Berlak, 1992; Gipps, 1994; Wiliam, 1994). These views of learning acknowledge that both students' existing knowledge and thinking processes influence the learning outcomes achieved and, therefore, both need to be taken into account in teaching (Bell, 1993). In taking into account students' thinking in their teaching, teachers are responding to and interacting with the students' thinking that they have elicited in the classroom. They are therefore undertaking formative assessment while teaching for conceptual development.

In science education, teaching for conceptual development arose from the 1980s' research on children's alternative conceptions (Driver, 1989). A central part of this teaching is dialogue (not a monologue) with students to clarify their existing ideas and to help them construct the scientifically accepted ideas (Scott, 1999). Therefore, giving feedback to students about how their existing conceptions relate to the scientifically accepted ones and helping them to modify their thinking accordingly, is both a part of formative assessment and teaching for conceptual development. Formative assessment is seen as a crucial component in teaching for conceptual development (Bell, 1995).

Formative Assessment and Learning

Due to these three trends in education, formative assessment is increasingly being used to refer only to assessment which provides feedback to students (and teachers) about the learning which is occurring during the teaching and learning, and not after. The feedback or dialogue is seen as an essential component of formative assessment interaction where the intention is to support learning (Clarke, 1995; Perrenoud, 1998; Sadler, 1989).

Assessment can be considered formative only if it results in action by the teacher and students to enhance student learning (Black, 1993). For example:

> The distinguishing characteristic of formative assessment is that the assessment information is used, by the teacher and pupils, to modify their work in order to make it more effective. (Black, 1995)

> Formative assessment has been defined as the process of appraising, judging or evaluating students' work or performance and using this to shape and improve students' competence. (Gipps, 1994)

It is through the teacher–student interactions during learning activities (Newman, Griffin, & Cole, 1989) that formative assessment is done and that students receive feedback on what they know, understand, and can do. It is also in these student–teacher interactions during learning activities that teachers and students are able to generate opportunities for furthering the students' understanding. As formative assessment is viewed as occurring within the interaction between the teacher and student(s), it is at the intersection of teaching and learning (Gipps, 1994). In this way, teaching, learning, and assessment are integrated in the curriculum.

Therefore, the process of formative assessment always includes students. It is a process through which they find out about their learning. The process involves them in recognizing, evaluating, and reacting to their own and/or others' evaluations of their learning. Students can reflect on their own learning or they may receive feedback from their peers or the teacher.

Formative assessment is also the component of teaching in which teachers find out about the effectiveness of the learning activities they are providing. It can be viewed as the process by which teachers gather assessment information about the students' learning and then respond to promote further learning. For example:

> Assessment should contribute to instruction and learning. . . . Assessment after instruction is over does not allow for the assessment to contribute to any instructional decisions. All that can be said is the degree to which a student mastered some amount of content. Assessment must be a continuous process that facilitates "on-line" instructional decision making in the classroom. (Gitomer & Duschl, 1995, p. 307)

Both formative and summative assessment influence learning. In other words, to improve learning outcomes, we need to consider not only the teaching and learning activities, but also the assessment tasks (Gipps & James, 1998). Moreover, the extent to which formative assessment improves learning outcomes is now being recognized. For example, Black and Wiliam (1998), in their review of classroom assessment, boldly state:

> The research reported here shows conclusively that formative assessment does improve learning. The gains in achievement appear to be quite considerable, and as noted earlier, amongst the largest ever reported for educational interventions. (p. 61)

While there has been much written on the importance of formative assessment to improve learning and standards of achievement (Harlen & James, 1996), there has been little research on the process of formative assessment itself. As Black and Wiliam (1998) suggest, there is a need to explore views of learning and their interrelationships with assessment.

The Learning in Science Project (Assessment)

This article reports on the findings of a research project investigating formative assessment in some science classrooms in New Zealand. This research, the Learning in Science Project (Assessment), is fully documented in Bell and Cowie (1997). This research was done under contract to the New Zealand Ministry of Education in 1995–1996 to investigate classroom-based assessment in science education in Years 7–10 (ages 11–14 years) in classrooms where the teacher of science was taking into account students' thinking (Bell, 1993). Four key aims for the research were:

1. To investigate the nature and purpose of the assessment activities in some science classrooms.
2. To investigate the use of the assessment information by the teacher and the students to improve the students' learning in science.
3. To investigate the teacher development of teachers with respect to classroom-based assessment, including formative assessment.
4. To develop a model to describe and explain the nature of the formative assessment process in science education.

Formative assessment in this research was defined as

> . . . the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning. (Cowie & Bell, 1996)

This definition parallels that of Sadler (1989, 1998), Gipps (1994), Clarke (1995), and Black and Wiliam (1998). The focus of this research was on formative assessment and not on assessment for qualifications, reporting to parents and caregivers, school-leaver documentation, or assessment for inspection or audit agencies. While these aspects of assessment in primary and secondary education cannot be separated entirely, they were not the focus of this research. Although continuous summative assessment was seen as formative assessment in the past, in this research, a distinction was made between the terms "formative assessment" and "continuous summative assessment."

The research was mainly qualitative, interpretive, collaborative, and guided by the ethics of care. Multiple data collection techniques were used, including interviews, surveys, and participant observation. The data documented in this article are illustrative rather than representative, given the constraints on space. Readers are referred to Bell and Cowie (1997) for a fuller documentation of the research methodology and data.

The research had three strands:

- *Ideas about assessment:* In this strand of the research, the views of assessment of nine teachers of science and one teacher of technology and some of their students were elicited at the beginning of the project and monitored throughout the project. Data for this strand was collected through interviews and surveys.
- *Classroom-based studies:* In this strand of the research, the classroom assessment activities, and in particular the formative assessment activities, of the ten teachers and

their students were studied and documented in the form of eight case studies (the first case study is the summary of the formative assessment of teachers 1, 4, and 6). Data for this strand was collected by participant observation; field notes; head notes; and documentary data, such as the writing on the board, student books, the wall displays, the teachers' "plan" for the unit, and the teachers' record books.

- *Teacher development studies:* In this strand of the research, teacher development activities were undertaken by the ten teachers to develop the formative assessment activities they used in their classrooms and to reflect on the data collected and analyzed. This occurred on eleven separate days over 1995–1996. Data for this strand was collected by audiotaped discussions, surveys, and field notes.

On one hand, the research was investigating the existing assessment practice of the ten teachers. On the other hand, the research was investigating the teachers' developing assessment practices over the two years of the project, while teaching to take into account students' thinking. In this way, it can be seen as investigating an intervention. This combining of research and development is more fully described in Bell and Cowie (1999).

The ten teachers who volunteered to take part in the research were primary (middle school) and secondary (junior high) teachers, women and men, beginning and experienced, and some had management responsibilities in the school. Some of the teachers had previous experiences of working on a research project in science education.

Each of the ten teachers chose a class to work with them on the project. For each teacher, this class of students changed in 1996, at the start of the new academic year. In total, there were 114 student interviews done during the course of the research.

One of the main findings was that the teachers involved were unable at the start of the research to explicitly describe what they did in the classroom that was called "formative assessment" (Bell & Cowie, 1997). This research has made explicit this often tacit process of formative assessment. In a previous article (Cowie & Bell, 1999), we outlined a model of formative assessment developed by the teachers and researchers involved in the research. It described two kinds of formative assessment: planned and interactive formative assessment. A key difference between the two was that of planning and purpose. In planned formative assessment (e.g., a prepared oral test of ten short questions at the beginning of the lesson), the teacher has planned the assessment before the lesson begins. In interactive formative assessment, the teacher is responsive to assessments that arise during his or her interactions with students in the lesson. While the teachers are prepared for interactive formative assessment (e.g., by increasing the opportunities for interaction with students during the lesson), they do not know the details of when and what assessment will occur. In this article, we wish to report on another key finding, that of the characteristics of formative assessment as described by the teachers during the teacher development days.

AN EXAMPLE OF INTERACTIVE FORMATIVE ASSESSMENT

A cameo which illustrates the process of interactive formative assessment follows (i.e., when the teacher notices, recognizes and responds to student learning and nonlearning in the classroom).

The lesson described was the first for the new term. Prior to the lesson, the teacher (teacher 7) stated she intended to review the techniques used for separating mixtures, to get the students to complete a written task, and to evaporate a salt and water solution. She began by telling the class that separating mixtures was the last topic in the unit they had been working on the previous term and that once this was completed, there would be a test.

The researcher field noted the lesson:

> The teacher introduced the topic of separating mixtures and commented that "some of these you have done before, so we should be able to rattle through it." She asked the students what they would do if she asked them to separate out the red smarties from a jar full of smarties. When the students stated they would pick them out on the basis of their color she linked their comments to separating: "the principle for separating components of mixtures . . . is to find something different and use that property to separate them."
>
> A textbook was distributed and the teacher read through and discussed the techniques of filtering, distillation, decanting, crystallising with the class . . . the teacher described the technique, sought ideas from the students, answered their questions and made links to their everyday experiences.
>
> Next, the teacher introduced a "thinking" task. She explained that she wanted the students to think about how they would separate out the two substances from each of the mixtures she was writing on the board.

How would you separate?	**technique**	**property**
kidney beans from broad beans		
oil from water		
iron filings from sand		
salt from sand		
dirt from water		
meths from water		
gold specks from sand		

> She discussed what techniques could be used to separate kidney and broad beans.

> The teacher moved around the class and spoke to a number of groups. She moved to the front of the class and said: "M has said 'What am I doing? I don't understand.' If she doesn't understand I am sure there will be lots of you who don't." The teacher read out "How do we separate kidney beans from broad beans?" A student asked what a kidney bean was. The students and teacher discussed the shape of kidney beans, the shape and color of both kidney beans and broad beans, and linked these features to how the beans could be separated.
>
> The teacher moved around the class talking to groups of students. . . . A group asked the teacher if they could use filtering to separate oil and water. The teacher went to the prep room and returned with oil and filter paper. She poured oil on the filter paper and discussed whether filtering was appropriate. She moved to the front of the class and demonstrated the effect of oil on filter paper, saying "If you are not sure of oil and filter paper . . ." While she was doing this another group asked about sand and salt. She collected these from the prep room and invited the students to come and look at them so they could "compare the size."
>
> She moved around the class talking to students. The teacher stopped that class and said: "A few things have become obvious . . . you need to know the properties of the different things and you need to know what happens when you put them together. . . . What about broad and kidney beans?"

Teacher 7, in an interview with the researcher later, highlighted the fact that her expectations were not realized. She had expected the students to be familiar with the techniques for separating substances, the properties of the substances she had selected for them to separate, and the way these substances reacted when they were mixed. The teacher spoke of being surprised that the students did not have a "general knowledge" and experience of the substances she was asking them to separate:

... they didn't know what kidney beans and broad beans were. ... I thought they were going to be really obvious things to use. But they didn't actually know what they were. ... it became obvious that they didn't actually know enough about these things to be able to separate them. ... there were quite a lot (of students) who didn't know what oil did. To me it was obvious it was going to float but it wasn't to them because they didn't have the experience of that. (T7/D1/96. That is, this is a quotation from teacher 7, made during classroom observation 1, in 1996)

When asked what formative assessment she considered that she had undertaken during the lesson, teacher 7 spoke of becoming aware of her use of formative assessment when she was introducing the task. Her awareness was triggered when she, unexpectedly, needed to ask a number of direct questions in order to obtain the information she needed. She said how she had become aware that the students had a limited knowledge of the properties of the substances she had included in the mixtures through their questions and comments. The teacher identified the students' lack of prior knowledge as a misconception that she had held. She said that she had expected that "they would know what salt and sand are like and be able to apply that theory." As a consequence of what she found out, she said " ... we needed to do something about the properties of salt and the properties of sand and then mix them and see" (T7/D1/96).

The actions the teacher took were to revisit the principles of separation, focusing on the use of a difference in the properties of the substances. She appeared to review the time frame for her goals. She stated her initial intention was to move quickly through the task and focus on separating salt and water through evaporation. On finding out the students' level of prior knowledge, her goal became one of increasing the students' knowledge of the individual properties of the substances she wanted them to separate and the relevance of these to the mixture. She began the next lesson by demonstrating the separation of oil and water, the students then separated meths and water and sand and iron filings.

The main points illustrated by this cameo are:

- Teacher 7 used interactive formative assessment (noticing, recognizing, and responding).
- Teacher 7 became aware of undertaking interactive formative assessment when the unexpected nature of the students' responses to the task necessitated her asking a number of questions. Questions and suggestions from the students while she was moving around the groups alerted her to the nature of the students' scientifically unacceptable ideas.
- The identity of the students who asked questions was significant. Some of the students who asked questions were among those she considered "thoughtful" and the most likely to understand (T7/D1/96).
- It was important that she was asked similar questions by more than one student, especially by more than one thoughtful student. This helped focus her attention and raised her awareness of the problems. In this instance, the student question about kidney and broad beans would have been sufficient to alert her to the students' ideas.
- Teacher 7 acted with both individuals and the class as a whole. For example, she showed a group the effect of oil on filter paper and then she demonstrated this to the whole class. This was a deliberate and considered action. She provided a number of reasons for this. She considered that the students who asked questions tended to be thoughtful, with a good understanding of ideas and said that if they were having problems, others would be too. She considered there were students who "do not like

to display their uncertainty to the teacher." (T7/D1/96) She said that some students who knew they could not do the task, could not formulate a question to ask her in order to obtain help. She acted with the whole class to provide feedback to all these students.

- Teacher 7 acted to address the students' scientifically unacceptable conceptions. She revisited the task requirements, the meaning of properties and techniques and she provided the materials for the students to look at. During the next lesson, she provided the students with the materials and they separated the mixtures themselves.

Further cameos and examples of formative assessment are given in Bell and Cowie (1997, 2000). When these instances of formative assessment were studied as part of the data analysis, nine key characteristics of formative assessment were identified.

KEY CHARACTERISTICS OF FORMATIVE ASSESSMENT

The nine characteristics of formative assessment discussed by the teachers were: responsiveness, the sources of evidence, the tacit process, using professional knowledge and experiences, an integral part of teaching and learning, formative assessment is done by both teachers and students, the purposes for formative assessment, the contextualized nature of the process, and the dilemmas (Bell & Cowie, 2001). Each of these will be discussed in turn, using illustrative data generated by the teachers during the teacher development days.

Responsiveness

The essence of formative assessment in the definitions cited earlier is the component of action or responsiveness of the teacher and students to the assessment information gathered or elicited. The different aspects of responsiveness discussed by the teachers are discussed in the following paragraphs.

Formative Assessment is Responsive in That It Is Ongoing and Progressive. The teachers involved in the research commented that they felt that formative assessment was characterized by its ongoing, dynamic, and progressive nature. They commented on the responsiveness:

> If you do something to find out where they (the students) are at, and then you do something from that to change your teaching or what you are doing, then its formative (assessment) . . . (TD5/96/14.13. That is, a quotation from the transcripts of teacher development day 5, in 1996, from audiotape 14, data segment 13)

Comments were made that formative assessment was not tied to a specific learning pathway and that the process was flexible and responsive:

> . . . there needs to be flexibility throughout your program . . . in order to take advantage of what . . . you've found. (TD4/95/11.39)

> A lot of the time you start off on one tack, and you think, no that didn't work so I'll try another tack, as so its self-assessment (of our teaching) as you go along. (TD4/95/11.41)

Formative assessment was seen as an ongoing, everyday event:

Without formative assessment, teachers do not function effectively. So it's your on-going, day-by-day, every-day assessment. (TD4/95/11.44)

Formative Assessment Is Responsive in That It Can Be Informal.   The teachers referred to assessment as both formal and informal. In saying this, they were usually referring to whether the information gathered was recorded and reported in some way or whether it was used in the classroom activities, without a written record being made. Formative assessment tended to be informal, with no written record of the information gathered. The information was used in the teaching and learning in the classroom and to build up a picture of the student learning by the teacher. For example:

It may just be how much concrete we set it in . . . I can go into a classroom and give the kids a spot ten-question test because I think they . . . just need to do that, to refocus them a bit . . . I don't record it anywhere, they'll do it in the back of their books. (We) mark it, and I say who got . . . this, who got that, 'that's fine,' and we carry on. And that's not set in any concrete at all. (TD4/95/11.74)

Formative Assessment Is Responsive in That It Is Interactive.   The teachers stated that formative assessment was interactive; that is, the information gathered was used in the interactions between teacher and student during the teaching and learning. For example:

A lot of people haven't been aware . . . that assessment can be done at other times . . . a lot of teachers . . . have just tended to assess students at the end of units and have really not been a part of that interactive process. (TD4/95/11.69)

Formative Assessment Is Responsive in That It Can Be Unplanned as Well as Planned.   The comments by the teachers suggested that at times they planned to do formative assessment, but at other times, they did unplanned assessments. A planned formative assessment was often used at the beginning of the unit; for example, the eliciting of students' prior knowledge before the teaching of the unit started. A planned formative assessment could also be used to start the formative assessment process within a lesson; for example, a quick ten-question spot test at the start of a lesson to find out if the students had understood the ideas introduced in the previous lesson.

The unplanned formative assessments arose from the students' responses, which often could not be predicted and planned for in advance. For example, in taking into account the student view that substances expand on cooling, teacher 5 responded by undertaking some unplanned formative assessment. The words "unanticipated" and "incidental" were also used in this context. For example:

I find that certainly toward the end of the year, children will ask these sorts of questions and so I planned for this to happen, but I never know what they are going to ask. (TD5/96/15.4)

The teachers commented that they planned or were prepared for the unplanned. For example:

Planned or unplanned . . . Yes, sure you get the kids set up. You don't know what you're going to get. And that's the unplanned part. What comes back from the children. But you get them set up in the first place. . . . So you plan the opportunity, but don't necessarily plan the (response). The lesson was planned, this is what they were going to do. But the unplanned

part was, oh . . . But that the most exciting teaching, when you sort of go tangent-wise. . . .
I know in that, I've ended up calling it planned and unplanned. I've now gone and changed
it to planned and incidental, which just sort of . . . 'cause unplanned makes it sound like
you don't know what you're doing, but . . . It is the planned opportunity, but there's also
that stuff that just opportunistic or spontaneous or some other word that I don't . . . Unan-
ticipated. . . . Ah, that's better. . . . Call it anticipated and unanticipated. (TD5/96/14.16;
TD5/96/14.17)

You have planned for the unplanned. I mean, you've left that opportunity for all those
incidental things that occur. (TD5/96/15.27)

Formative Assessment Is Responsive in That It Can Be Proactive or Reactive.
The terms proactive and reactive were also used to indicate the notion of responsiveness
inherent in formative assessment. That is, the teacher could be proactive in deliberately
seeking formative assessment information from students or reactive, when they undertook
formative assessment in response to other information they had gathered about the students'
learning. For example:

I thought, it could be proactive where you actually go out and you seek, um, specific times
throughout a lesson to actually do the formative assessment. Or it could be reactive. I find
that a lot of teaching is, a great percentage is, reactive teaching. (TD5/96/15.12) . . . (for
example, a) crisis, where students, for some reason, it may be that they are off task, or not
prepared or . . . inattentive at listening, or they're all dependent on being followers, lack of
ideas or just lost. Crisis point where formative assessment comes in, you have to sort of step
in there and take a real lead. (TD5/96/15.14) . . . And then, there was a refocusing because
you get students who tend to go off track. So by asking questions, on a fairly informal basis,
you find out that this kid is way off track and really not going to achieve the objectives that
I had planned for the unit. So then you have to get them to refocus again. Like exploring
alternative methods and backtrack, taking them back. (TD5/96/15.15)

***Formative Assessment Is Responsive in That It Involves Responding with Indivi-
duals and with the Whole Class.***   The teachers described the way they moved back and
forth between the whole class, a group, and an individual in their interactions as a result of
gathering formative assessment information. For example:

But there is an interesting issue that's coming up for us, and that is that the interplay between
the child or the student and the class and how information about the general class feeds
into what we do with the student and how information we find out about from a particular
student can then feed back into the whole class. That's where the interaction between those
two. . . . Looking at the class or looking at the child are all related. (TD5/96/14.26)

Formative Assessment Involves Uncertainty and Risk Taking.   As the formative as-
sessment done by the teachers was often unplanned and responsive, it involved uncertainties
and taking risks. Formative assessment involved the teacher finding out and responding to
the diverse views of students; it had indeterminate outcomes; it could not be planned in detail
before the lesson; the effects of the required actions were not usually known beforehand;
and usually it required the teacher to take action in the busy-ness of the classroom. Their
confidence in their professional knowledge and skills was seen by the teachers to influence
the degree of risk and uncertainty taken.

Formative Assessment Has Degrees of Responsiveness.   The teachers said that they had to manage the degree of responsiveness when doing formative assessment. They were aware that they had to manage the behavior and learning of the whole class as well as that of individuals. They also had to manage attending to the students investigating their own interests and ideas, and to the students learning what was listed in the curriculum. In both these situations, responding to one aspect meant that they could not respond to the other. They could not always be as responsive to a situation as they wished or were able to.

In summary, the teachers felt that a characteristic of formative assessment was that it was responsive. This responsiveness was discussed in terms of formative assessment being ongoing; dynamic and progressive; informal; interactive; unplanned as well as planned; reactive as well as proactive; with the class, group, or individual; involving risk and uncertainty; and managing the degree of responsiveness.

The Sources of Information and Evidence

The second characteristic of formative assessment was the sources of information and evidence. Formative assessment, like summative assessment, may use student written or oral work, but the teachers commented that formative assessment relies on nonverbal as well as verbal information. For example:

> (teachers will be) . . . observing kids, in terms of . . . facial expressions, body language, listening, talking, writing . . . (TD4/95/12.7)

The sources of formative assessment information for the teachers included the teachers' observations of the students working (e.g., in practical activities; the teachers reading student-written work in their books, posters, charts, and notes; and the teachers listening to students' speech, including their existing ideas, questions and concerns, and the new understandings they were developing). The teachers set up different learning situations to provide the opportunities for this information to be gathered or elicited. For example, the teachers organized practical and investigative work, brainstorming, spot tests, students recording their before-views, library projects, watching a video, whole class discussions, and student self-assessment activities. There was acknowledgment that different learning situations enabled different formative assessment information to be elicited.

A Tacit Process

A third characteristic of formative assessment is that it is often a tacit process. A frequent comment from the teachers was that they were not always consciously aware of doing formative assessment, and in particular unplanned or interactive formative assessment. For example:

> I am still not recognizing what I do in terms of unplanned (interactive formative assessment). (TD5/96/14.1)

The teachers tacitly undertook formative assessment and were not always able to explicitly describe it to the researchers. This unawareness was evident in the discussion by the teachers of using "gut feelings." For example:

> Because you can stay awhile with a group . . . oh, I'll just listen to the kids. And that's where you get your gut assessment. (TD5/96/15.24; TD5/96/15.25)

And you don't get a gut feeling sitting the night before thinking about . . . It's when you're there, so you are interpreting something that's happening in the room . . . That's what your gut feeling is. . . . How many times do you actually change your teaching style or whatever, during the lesson, because the gut feeling gives. . . . You've said something and you know exactly where you want to go with the kids, and something sort of happens, and it's not working, so you sort of get that gut feeling. You think pretty fast then. . . . But you can't tell other people that you work by gut feelings, because they need something tangible that they can actually . . . think about and something you can rationalize. You can't rationalize just in cold turkey gut feelings. (TD7/96/20.39)

The teachers also spoke of "getting an impression" of the class. For example:

And the formative assessments . . . could be . . . more formal tasks or they could be just impressions in the classroom . . . we can't really identify what tells us that the majority of the class know the first bit so we can go onto the second bit, but to me, that automatic assessment is part of formative assessment. (TD4/95/11.30; TD4/95/11.31)

The experiences of being involved in the research had made more visible to the teachers what formative assessment information they were collecting and what they were doing with it. For example:

I personally never really realized I was doing it (formative assessment) except that the class was with me, or not with me. And since this experience, you sort of tend to focus more on what am I actually taking in here. Or what is it actually telling me . . . this process is going on. And I think for most teachers it will still be subtle and not obvious. (TD9/96/27.3)

The teachers stated that thinking about formative assessment had helped them become more aware of their professional knowledge and skills and more able to use these in the formative assessment process in the classroom.

Using Professional Knowledge and Experiences

A fourth characteristic of formative assessment is that it was seen to rely on a teacher's professional knowledge and experiences. The professional knowledge and experiences of the teachers were seen as important in attending to some sources of information (rather than others), in interpreting the elicited information, and in taking action. This professional knowledge and experience included the teachers' knowledge and experiences of the topic, of the students as learners, and from having taught the unit of work before. For example:

Because the knowledge of how to teach is what makes that successful. . . . you've got to have confidence in your ability to teach.

And that's all your other skills . . . those you actually can't do without.

. . . Yeah, it's knowing how to handle the students, isn't it?

I still think that the knowledge base of the subject has got a place, though. I mean,

. . . I think, too, if your knowledge is at a reasonable level, you can take advantage of the one off situations that sometimes happen. Whereas if it's not there, you can't take advantage at all.

And the more you teach, the better you become. (TD9/96/26.41)

An Integral Part of Teaching and Learning

Another characteristic of formative assessment was the action taken by the teacher and the student as a result of the information gathered. Taking action to enhance learning is an integral part of the definition of formative assessment. The teachers commented on the variety of actions they took in response to the formative assessment information as well as the way in which they evaluated their actions. For example:

> I was thinking the teacher would get some information, interpret it, decide to act, they would act, and then it may or may not work, and they would react to that. . . . Deciding on a new experiment, deciding to do a discussion, what sort of teaching reaction. . . . To make a decision where you go. Are you going to reteach it or are you going to have a look at it from a different (perspective), go get some more information. . . . Move on to the next step, cause they've got it. . . . you interpret information see, then you act on, then you react. (TD7/96/20.44)

Taking the action involved the teachers making decisions and judgments, using their professional knowledge and experiences. The action often appeared to the teachers to be a part of teaching and the comment was made as to whether the action was a part of teaching or a part of assessment. The overlap between the action inherent in formative assessment and teaching was frequently acknowledged. For example:

> I think formative assessment and teaching . . . overlap really. (TD5/96/14.9)

The teachers described their actions as those to facilitate students' learning. They spoke of actions that mediated the students' learning of the science and actions that enhanced the personal and social development of the students. The actions taken were, for example, suggesting further questions, suggesting further activities, questioning of a student's ideas, explaining the science, giving feedback as to the students scientifically acceptable or unacceptable ideas. The notion of the teacher as a neutral facilitator was not seen as part of formative assessment:

> Being a neutral facilitator isn't what we're on about here. In terms of formative assessment you (are) wanting to take action, you may choose to do nothing because you want to leave the kids for a while to see if they can find their way through it, but if they can't, you might want to then make another decision. (TD7/96/20.67)

A similar comment could be made as to whether the action taken by the students was a part of learning or assessment. These comments highlighted that formative assessment is an integral part of teaching and learning.

Formative Assessment Is Done by Both Teachers and Students

Another characteristic of formative assessment was that both the teacher and the student were doing the assessing. The teachers' comments highlighted the involvement of students as assessors, in addition to the teachers. Part of a discussion of the model of formative assessment, and, in particular, the cycle of gathering information, interpreting, and taking action is following:

> Think of it from the kid's point of view, the kid gathers information from what you've given them already, they filter it, decide what's relevant to them, they interpret what they need to do however they like, they act on that information, and then from whatever you do or from whatever things happen, they gather more information and so on.

So it works exactly for them. It's just that our acting becomes their gathering information points.

Whatever we do . . . they get the information from (it).

And their acting is our gathering. (TD7/96/20.45; TD7/96/20.59)

Student self-assessment was seen as an important part of formative assessment:

Formative assessment isn't just for the teacher, it's for the students to know that they are still moving, and going somewhere . . . So it's a decision making process for the student. (TD4/95/11.38)

Some teachers used the phrase "self-assessment" to refer to the teachers evaluating their own teaching:

I think it's self-assessment by the teacher as they go along. As well as the other side which is helping the students assess themselves . . . I think it's what we automatically do—assessing ourselves as we go along . . . and the kids . . . assess themselves. (TD4/95/11.40)

The Purposes for Formative Assessment

The seventh characteristic of formative assessment was the purposes for which it was done. As documented in the case studies, the teacher development days and the surveys (see Bell & Cowie, 1997), the teachers identified several purposes for formative assessment. In particular, they identified that the two main purposes of formative assessment were to inform the students' learning and to inform their teaching.

The purposes to support the students' learning included monitoring the progress, learning, or understandings of the students during the teaching and learning. For example:

. . . The teachers' purpose for the thing that they're doing at the moment, what they want the children to learn, what they are trying to get out of them, the kind of thing that they do. (TD7/96/20.31)

The learning might be social, personal, or science learning (Bell & Cowie, 1997):

So in that purposes for learning, . . . you do have science purposes . . . You did have a social purpose and a personal purpose. (TD7/96/20.55)

The purposes to support learning also included giving feedback to students about what learning was valued in the classroom, giving legitimacy to the students' scientifically acceptable ideas, supporting long- or short-term goals, and finding out whether an activity or task was "working." For example:

. . . Is there a case when you just check to see whether the activity is working or not? Like, I mean, just thinking from my own teaching where you might set up a group activity and you realize after visiting two or three groups that the instructions haven't been clear enough, so you stop the class and say, look instruction number 3, I've actually missed a bit out. It should read like this, and everyone nods and away you go. So in this case you haven't actually checked that, you haven't gone right back to the learning goals, you've just got to the level of "is this activity working or not."

I was just thinking on that train of thought too because it's sort of as if we're a trouble shooter and just watching if things are moving in the right direction . . . You have to intervene in some way. It could be to the whole group or an individual.

But how do you do it. I suppose, like we said, do you stop the group or do you speak to the individual on the side. . . . (TD7/96/20.71)

The purposes to support teaching (mentioned by the teachers) included planning in the current lesson and unit; planning for future teaching; knowing when to input new ideas and when to move on to the next topic; knowing when to introduce an activity to maintain interest and motivation; evaluating the actions taken in previous formative assessments and teaching activities; finding out if the students had understood or not; providing information to report to students, caregivers, and the school; and providing assessment information additional to the quantitative marks on achievement in reporting.

The Contextualized Nature of Formative Assessment

The eighth characteristic was that the formative assessment undertaken by the teachers was contextualized. In other words, the purposes, the information elicited, the interpretations made, the actions taken, depended on many contextual factors. For example, the ways the formative assessment information was elicited, interpreted, and acted on was influenced by the learning situations used (whole class, small groups, or individuals); by the learning activities chosen (e.g., brainstorms, investigations, watching a video, and library projects); the teacher's knowledge of the students; the professional knowledge and skills of the teacher; the topic of the lesson and the teacher's purposes for the lesson. This characteristic has suggested that formative assessment may be seen as a sociocultural and discursive activity (Bell, 2000; Bell & Cowie, 2001).

Dilemmas

The ninth characteristic of formative assessment was that of the dilemmas faced by the teachers when doing formative assessment. The interaction between these characteristics in the processes of formative assessment presented the teachers with dilemmas. The word "dilemmas" is used as there was no obvious solution to the situation and the decision made in response to each situation would depend on contextual features and the teacher and students concerned. Unlike problems that can be solved, dilemmas are managed and this management relies heavily on the professional judgment of teachers. The nature of these dilemmas was evident in the discussions on the teacher development days on the tensions between formatively assessing the class or an individual; between formatively assessing the science or the personal and social development; between formatively assessing the science in the curriculum and the science outside the curriculum; and between the different purposes for eliciting and taking action.

FINAL COMMENTS

A tenth characteristic of formative assessment was clarified from the data generated by the students (Cowie, 2000), the characteristic of student disclosure. Disclosure relates to the extent to which a task or activity produces evidence of student performance or thinking. In the classrooms, the teachers used tasks and strategies to elicit student ideas and

students voluntarily disclosed their ideas by asking questions and discussing their ideas, but the students did not automatically and fully disclose their thinking. In discussing the extent to which they disclosed their ideas to the teacher, the students commented on the nature of the assessment strategies used by the teachers, the relationship between teachers' rights and disclosure, disclosure as a source of potential harm, and trust as mediating the disclosure (Bell & Cowie, 2001). An important finding of this research was that the validity of formative assessment relied on the extent of student disclosure (Cowie, 2000).

In summary, the nine characteristics of formative assessment that were identified by the teachers were responsiveness; the sources of evidence; the tacit process; using professional knowledge and experiences; an integral part of teaching and learning; formative assessment is done by both teachers and students; the purposes for formative assessment; the contextualized nature of the process; and dilemmas. The value of documenting the characteristics of formative assessment to the teachers in the research was that a largely tacit process was made more visible and explicit for teachers wanting to improve their practice in either pre-service or in-service teacher education.

REFERENCES

Bell, B. (1993).Taking into account students' thinking: A teacher development guide. Hamilton, New Zealand: University of Waikato.

Bell, B. (1995). Interviewing: A technique for assessing science knowledge. In S. Glynn & R. Duit (Eds.), Learning Science in Schools: research reforming practice. New Jersey: Lawrence Erlbaum Associates.

Bell, B. (2000). Formative assessment and science education: Modelling and theorising. In R. Millar, J. Leach, & J. Osborne (Eds.), Improving science education: The contribution of research. Buckingham: Open University Press.

Bell, B., & Cowie, B. (1997). Formative assessment and science education. Research report of the Learning in Science Project (Assessment). Hamilton, New Zealand: University of Waikato.

Bell, B., & Cowie, B. (1999). Researching formative assessment. In J. Loughran (Ed.), Researching teaching. London: Falmer Press.

Bell, B., & Cowie, B. (2001). Formative assessment and science education. Dordrecht: Kluwer Academic Publishers.

Berlak, H. (1992). The need for new science of assessment. In H. Berlak, F. M. Newmann, E. Adams, D. A. Archbald, T. Burgess, J. Raven, & T. A. Romberg (Eds.), Toward a new science of educational testing and assessment (pp. 1–22). Albany: State University of New York Press.

Black, P. (1993). Formative and summative assessment by teachers'. Studies in Science Education, 21, 49–97.

Black, P. (1995). Can teachers use assessment to improve learning? British Journal of Curriculum & Assessment, 5(2), 7–11.

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

Brown, S. (1996). Summary Comment. Symposium on Validity in Educational Assessment. Educational Assessment Research Unit, University of Otago, Dunedin, New Zealand, 28–30 June.

Clarke, D. (1995). Constructive assessment: Mathematics and the student. In A. Richardson (Ed.), Flair: AAMT Proceedings. Adelaide: AAMT.

Cowie, B. (2000). Assessment in science classrooms. Unpublished DPhil thesis. Hamilton, New Zealand: University of Waikato.

Cowie, B., & Bell, B. (1996). Validity and formative assessment in the science classroom. Invited keynote paper to Symposium on Validity in Educational Assessment, 28–30 June, Dunedin, New Zealand.

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. Assessment in Education, 6(1), 102–116.

Driver, R. (1989). Students' conceptions and the learning of science. International Journal of Science Education, 11(5), 481–490.

Duschl, R., & Gitomer, D. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. Educational Assessment, 4, 37–73.

Erickson, G., & Meyer, K. (1998). Performance assessment tasks in science: What are they measuring? In B. Fraser & K. Tobin (Eds.), International handbook of science education. Dordrecht: Kluwer Academic Publishers.

Gardner, H. (1985). Frames of mind: The theory of multiple intelligences. USA: Basic Books.

Gipps, C. (1994). Beyond testing: Towards a theory of educational assessment. London: The Falmer Press.

Gipps, C., & James, M. (1998). Broadening the basis of assessment to prevent the narrowing of learning. The Curriculum Journal, 9(3), 285–297.

Gitomer, D., & Duschl, R. (1995). Moving towards a portfolio culture in science education. In S. Glynn & R. Duit (Eds.), Learning science in schools: Research reforming practice. New Jersey: Lawrence Erlbaum Associates.

Gitomer, D., & Duschl, R. (1998). Emerging issues and practices in science assessment. In B. Fraser & K. Tobin (Eds.), International handbook of science education. Dordrecht: Kluwer Academic Publishers.

Harlen, W., & James, M. (1996). Creating a positive impact of assessment on learning. New York: AERA.

Johnson, S. (1989). National assessment: The APU approach. London: Her Majesty's Stationery Office.

Keeves, J., & Alagumalai, S. (1998). Advances in measurement in science education. In B. Fraser & K. Tobin (Eds.), International handbook of science education. Great Britain: Kluwer Academic Publishers.

Ministry of Education (1993). The New Zealand Curriculum Framework. Wellington: Learning Media.

Ministry of Education (1994). Assessment: Policy to Practice. Wellington: Learning Media.

National Research Council (1999). The assessment of science meets the science of assessment. Board on Testing and Assessment Commission on Behavioral and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press.

Newman, D.,Griffin, P., & Cole, M. (1989). The construction zone: Working for cognitive change in school. Cambridge: Cambridge University Press.

Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. Assessment in Education, 5(1), 85–102.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144.

Sadler, R. (1998). Formative assessment: Revisiting the territory. Assessment in Education, 5(1), 77–84.

Scott, P. (1999). An analysis of science classroom talk in terms of the authoritative and dialogic nature of the discourse. Paper presented to the 1999 NARST Annual Meeting. Boston, MA.

Tamir, P. (1998). Assessment and evaluation in science education: Opportunities to learn and outcomes. In B. Fraser & K. Tobin (Eds.), International handbook of science education. Dordrecht: Kluwer Academic Publishers.

Wiliam, D. (1994). Towards a philosophy for educational assessment. Paper presented to the Annual Conference of the British Education Research Association, Oxford.