

Formative and Summative Interpretations of Assessment Information

John Hattie

School of Education

The University of Auckland

Auckland, New Zealand

January 2003

## FORMATIVE AND SUMMATIVE INTERPRETATIONS OF ASSESSMENT INFORMATION

There have been many articles claiming that formative rather than summative assessment needs to be the priority of teachers. The typical claim is that formative assessment is about discovering *what* the learner knows, understands or can do. Summative assessment, in contrast is about *whether* the learner knows, understand or can do a predetermined thing (cf., Torrance & Pryor, 1998, p. 153). Hence, the conclusion that teachers need more formative assessment and less summative assessment. The aim of this paper is to dispel this claim, especially as the repetition of this claim has major negative implications for the quality of teaching and learning.

We first set the scene by claiming that assessment is about feedback information (and not about “tests” although tests can be one mechanism to provide feedback information)

### **Setting the Scene: Assessment Is Feedback Information**

Assessment is primarily concerned with providing teachers and/or students’ feedback information, which they then need to interpret when answering the three feedback questions: Where am I going?, How am I going? and Where to next?. Specifically, feedback is actions or information provided by an agent (e.g., teacher, peer, book, parent, experience) that provides information regarding aspects of one’s performance or understanding. Feedback is information about the task that fills a gap between what is understood (How am I going?) and what is aimed to be understood (Where am I going?). It can lead to increased effort, motivation or engagement to reduce the discrepancy between current status and the goal, it can lead to alternative strategies to understand the material, it confirm to the student that they are correct or incorrect, it can indicate that

more information is available or needed, it can point to directions that the students could pursue, and it can lead to restructuring understandings (Where to next?).

Clearly assessment needs to relate to providing answers to the three feedback questions. Thus, teachers and students need to devise assessments and seek information about the task, about the processes or strategies to understand the tasks, about the regulation, engagement and confidence to become more committed and learn, and/or about the discrepancy between current status and the goal. As Shephard (2000) noted, the successful teacher is “able to ask the right questions at the right time, anticipate conceptual pitfalls, and have at the ready a repertoire of tasks that will help students take the next steps requires deep knowledge of subject matter” (p. 12).

Crooks (1988), and more recently, Black and Wiliam (1998) have demonstrated there is little evidence that classroom assessment has assisted in the learning process. Black and Wiliam, for example, reviewed 578 publications relating to the role of assessment in learning and concluded that classroom assessment “typically encourages superficial and rote learning, concentrating on recall of isolated details, usually items of knowledge which pupils soon forget, ... teachers do not generally review the assessment questions that they use and do not discuss them critically with peers, so there is little reflection on what is being assessed” (p. 17). Black and Wiliam highlighted the potential of assessment to provide feedback, and the importance of the social context of learning. We have argued elsewhere (Hattie & Jaeger, 1998) that Black and Wiliam needed to go further and should have argued that effective learning and assessment occurs not by being concerned about the testing itself, but by the manner in which the assessment is integrated into the cycle of challenging goals—student receptivity to the challenges—and feedback.

This cycle requires excellent teachers who are thinking, reflective, enthusiastic, passionate, and knowledgeable about the content. These teachers need to care to make the difference, preferably caring about their subject matter, and caring about all of their students' achievements. Excellent teachers encourage their students to share in seeking answers to our three feedback questions.

There are many ways in which teachers can deliver feedback to students, and for students to receive feedback from teachers, peers and other sources. Tests can perform this function but too often they are devoid of much feedback to the students and thus can be an inefficient method of providing feedback.

### **Assessment and Formative and Summative Interpretations**

The terms formative and summative were introduced by Michael Scriven 30+ years ago (Scriven, 1967), and it is worth tracing the history of these terms and their various uses and abuses (Scriven, 1967). As Scriven outlined, the distinction between normative and summative is more related to interpretations and time—as illustrated by Bob Stake's maxim: When the cook tastes the soup it is formative, when the guests taste the soup it is summative. Thus, a key issue is timing, and it is possible that the same stimulus (e.g., tasting the soup) can be interpreted and used for both forms of assessment. Hence, it is NOT the instrument (tasting) that is formative and summative it is the timing of the interpretation and the purpose to which the information is used.

Similarly, it is possible to take what many consider a “summative test”, the PATs or asTTle and use the feedback information from them to make decisions to enhance the program (improve the soup) while in the making, or to make statements about students when they have completed a program of study. Hence, it is the interpretations and NOT

the test/form of assessment that is formative or summative. The same interpretations could be used both in a formative and summative manner. As Scriven has noted it is a fallacy to assume that formative and summative represent two types of assessment. Instead, they refer to interpretations of information at two differing times—either to make interpretations that can then lead to changing a program of learning, or that can lead to a statement about the learning at the end of the program.

In the same way that the goal of the cook is to make the best soup possible for the guests, it is imperative that teachers have excellence summative evaluation in place in their classes (and Principals in their school) as that can be among the more powerful evidence that there is likely to be excellent formative evaluation in place. Poor soup for the guests is pretty powerful evidence of poor cooking. If a school has poor summative assessment in place then it is unlikely they will have the ability, purpose or wherewithal to be concerned with formative assessment. Serving poor soup to the guests is probably the best indicator that the cook was lousy at tasting it during the preparations. Too much reliance on tasting the soup, may lead to inattention to the goals—such as making soup cold for when the guests arrive. The goals or learning intentions are critical aspects of this process, and can have major effects on the usefulness of the formative and/or summative evaluation.

It is worth rehearsing some of the further claims Scriven has made about formative and summative assessment. First, sometimes we do not know when administering an assessment whether it is formative or summative. For example, a teacher may consider evaluating the course of teaching by using an end-of-course test, but then find that remedial work is necessary making the test interpretations more formative than intended.

Too often formative interpretations are used to make teachers believe that they have made progress, and thus time to move on to other topics etc., hence formative interpretations can be turned into summative interpretations. Too often, however, an emphasis on “improvement” as if “any” evidence of improvement is necessarily a sign of progress is a weak yardstick indeed. When over 500,000 studies assessing the influence of many aspects to enhance achievement (e.g., policy, student prior learning, home, school, curricula, teacher), over 95% of all effects indicate students have “improved”. Almost everything works. So saying that there is evidence of “improvement”/ progress is trivial—such evidence is rarely and barely enough. It is the level and rate of progress that is critical, and it must be related to some defensible goal of progress (in effect-size terms at least .40). Too often in education, the goal posts are set so low (at anything greater than zero), too often satisfaction that there is progress ( $> 0$ ) is the goal, and we do not question vigorously enough whether *sufficient* progress has been made, or whether the progress goal was high enough.

That the distinction between formative and summative relates more to the timing of the interpretations, than to the form of assessment has major implications for the quality of the interpretations from both. Undertaking formative evaluation any less rigorously than summative evaluation undermines the accuracy of the mid-course corrections, which is all too likely to send the mission in the wrong direction. Contrary to popular utterance, it is the formative assessments that need to be more rigorous, as too often mid-course corrections, evaluations about progress are based on very weak evidence, and very low goal posts. Too often, we underestimate the importance of accuracy and dependability in formative assessment—diagnosis and corrections need to be excellent.

## Conclusions

Assessment is primarily concerned with providing teachers and/or students' feedback information, and thus needs to address three questions: What are the learning intentions? How am I going? And Where to next?

The distinction between normative and summative is more related to interpretations and time as illustrated by Bob Stake's maxim: When the cook tastes the soup it is formative, when the guests taste the soup it is summative. It is NOT the instrument that is formative and summative it is the timing of the feedback. Sometimes we do NOT know when we administered an assessment whether it is formative or summative. Summative evaluations can be powerful in education and Formative evaluations can be ineffectual in education (and vice versa).

Too often formative is seen as less rigorous or more informal than summative, when it needs to be the opposite. Doing formative evaluation any less rigorously than summative evaluation undermines the accuracy of the mid-course corrections, which is all too likely to send the mission in the wrong direction.

Schools that create a climate that all are responsible for the progress of the students, schools that deprivatises the information and evidence, and schools that collaborate to improve learning are great schools—it is simple. When did you meet with colleagues and talk about the evidence of progress of your and their students, talk about how to improve your teaching, how to change your teaching, and how to do this in light of the evidence that what you are currently doing is just not good enough to have the effect on progress of students? Do you feel psychological safe to discuss how to improve your teaching (not the kids, not the curriculum, not the resources, not the class size, not the conditions) but

YOUR teaching, with colleagues. Such schools are assessment rich schools where the probability of great learning is occurring, and the teachers are mirroring that they ask of their students—let’s learn together and respect each other by seeking evidence we are doing a great job. Teachers need to share assessment, students, and teaching. Teachers need to see assessment as saying something about what they (not the kids) are doing.

We need to use the terms formative and summative to relate to the timing of interpretations, and not as if they pertain to a type of test. We need to talk about the quality of the evidence and the nature of the interpretation, and the excellence of the decision making.

Who says feedback about the quality of my teaching matters—me. When you use asTTle, Exemplars, ARBs, NEMP or any assessment tool, I hope you do learn about your teaching while you are seeking evidence about your students progress and about what is not working and thence to improve the teaching and learning, and that this informs you about the closeness and adequacy of your learning intentions, and that it does help in deciding where to go next. The art and science of effective student learning is directly related to the quality of decision making by the teacher—go forth and seek that evidence in dependable formative and summative ways.



### References

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-481.
- Hattie, J., & Jaeger, R. (1998). Assessment and classroom learning: A deductive approach. *Assessment in Education: Principles, Policy & Practice*, 5(1), 111-122.
- Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), *Perspectives of curriculum evaluation* (Vol. 1, pp. 39-55). Chicago: Rand McNally.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Florence, KY: Taylor & Francis.