

# Developing the theory of formative assessment

Paul Black · Dylan Wiliam

Received: 18 September 2008 / Accepted: 30 December 2008 /  
Published online: 23 January 2009  
© Springer Science + Business Media, LLC 2009

**Abstract** Whilst many definitions of formative assessment have been offered, there is no clear rationale to define and delimit it within broader theories of pedagogy. This paper aims to offer such a rationale, within a framework which can also unify the diverse set of practices which have been described as formative. The analysis is used to relate formative assessment both to other pedagogic initiatives, notably cognitive acceleration and dynamic assessment, and to some of the existing literature on models of self-regulated learning and on classroom discourse. This framework should indicate potentially fruitful lines for further enquiry, whilst at the same time opening up new ways of helping teachers to implement formative practices more effectively.

**Keywords** Formative assessment · Assessment for learning · Dynamic assessment · Self-regulation · Instruction · Pedagogy · Cognitive acceleration · Dialogue

## 1 Introduction

The purpose of this paper is to develop the theory of formative assessment beyond the stage reached in our earlier writing, drawing on a variety of sources in the literature that have addressed this issue, whether directly or obliquely. Our earliest work on formative assessment (Black and Wiliam 1998a, b) did not start from any pre-defined theoretical base but instead drew together a wide range of research findings relevant to the notion of formative assessment. Work with teachers to explore the practical applications of lessons distilled therefrom (Black et al. 2002,

---

P. Black (✉)  
Department of Education and Professional Studies, King's College London,  
Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, UK  
e-mail: paul.black@kcl.ac.uk

D. Wiliam  
Institute of Education, University of London, London, UK

2003) led to a set of advisory practices that were presented on a pragmatic basis, with a nascent but only vaguely outlined underlying unity. So our *first* aim in this paper is to provide a unifying basis for the diverse practices which are said to be formative.

Subsequently, (Black and Wiliam 2006) we explored the changes that occurred in the classrooms of teachers developing formative assessment, and proposed a theoretical frame for the study of such classrooms. However, this theoretical frame was grounded in the data collected from classroom observations and interviews with teachers, and no systematic attempt was made to connect these data to work on such topics as classroom practice, or the regulation of learning. Other recent work has focused on aspects of implementation, notably on effecting change with communities of teachers (Wiliam 2007a) and on problems of superficial adoption (Black 2007), whilst both the book by Black et al. (2003) and the studies of the project on “Learning how to learn” (James et al. 2007) have discussed the learning principles underlying formative practices

In the conclusion of our 2006 article, we raised the wider issue of the role of formative assessment:

Thus, whilst we cannot argue that development of formative assessment is the only way, or even the best way, to open up a broader range of desirable changes in classroom learning, we can see that it may be peculiarly effective, in part because the quality of interactive feedback is a critical feature in determining the quality of learning activity, and is therefore a central feature of pedagogy. (p. 100)

This introduces our *second* aim in this paper, which is to locate formative interactions within more comprehensive theories of pedagogy.

Perrenoud (1998), commenting on our 1998 review, further emphasised the need to place any treatment in a broader context of studies of formative assessment:

This [feedback] no longer seems to me, however, to be the central issue. It would seem more important to concentrate on the theoretical models of learning and its regulation and their implementation. These constitute the real systems of thought and action, in which feedback is only one element. (p. 86)

This expanded the agenda, for the issues that it raised require that a wider range of theories be considered, so that the concept of formative interaction may be enriched and contextualised in the light of relevant theories. This task of linking our analysis to other theoretical writing about learning interactions will be our *third* aim.

Finally, the argument must return to its starting point in the classroom practices of formative assessment, so we shall consider, as a *fourth* aim, how our findings, and any further theoretical reflection, might suggests ways to extend and/or improve those practices.

Our analysis will be exploratory rather than definitive: its main value might be to guide development of further enquiries, so our *fifth* aim will be to offer suggestions about such enquiries.

In what follows, we shall first set out in Section 2 a summary of recent work that addresses our first aim. Then in Section 3 we shall suggest a general model of teacher-learner interactions that will set the scene for consideration, in Section 4, of

how teachers might interpret the responses of learners. The next step is to consider how teachers might frame and steer their feedback to any responses. This is discussed at two levels. In Section 5 we discuss the strategic level, i.e. how the formulation of feedback is guided by the teacher's broader pedagogical orientation. In Section 6 we then discuss the tactical level, looking at the fine-grain of types of response and the lessons learned from studies of classroom discourse.

There is not a one-to-one match between these sections and our five aims. We shall examine within each section how it contributes to particular aims. This will be the basis for Section 7 which will present our overview of the extent to which our treatment has met our aims.

## 2 A unifying basis

The purpose of this section is to draw together ideas developed in several earlier publications in order to serve our first aim of helping to provide a unifying basis for the diverse practices that are said to be formative.

Early work on formative assessment centred on five main types of activity, suggested by evidence of their potential effectiveness, and developed with and by teachers in normal classroom work (Wiliam 2000; Black et al. 2003; Wiliam 2007b):

- Sharing success criteria with learners
- Classroom questioning
- Comment-only marking
- Peer- and self-assessment
- Formative use of summative tests

However, whilst each of these five broad headings appeared to be connected to the central idea of formative assessment, precisely how they are so connected was not clearly articulated. Furthermore, the lack of a theoretical foundation raised questions about whether these five collectively exhaust the domain of formative assessment practice.

In order to provide a better theoretical grounding for formative assessment, Wiliam and Thompson (2007) drew on Ramaprasad's (1983) three key processes in learning and teaching:

- Establishing where the learners are in their learning
- Establishing where they are going
- Establishing what needs to be done to get them there

Traditionally, the teacher has been regarded as responsible for each of these three, but it is also necessary to take account of the role that the learners themselves, and their peers, play in them. The teacher is responsible for designing and implementing an effective learning environment, and the learner is responsible for the learning within that environment. Furthermore, since the responsibility for learning rests with *both* the teacher *and* the learner, it is incumbent on each to do all they can to mitigate the impact of any failures of the other (in the language of partnership law, teachers and learners are jointly and severally liable!).

|         | Where the learner is going   | Where the learner is right now  | How to get there  |
|---------|--|---|---|
| Teacher | <b>1</b> Clarifying learning intentions and criteria for success       | <b>2</b> Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding | <b>3</b> Providing feedback that moves learners forward |
| Peer    | Understanding and sharing learning intentions and criteria for success | <b>4</b> Activating students as instructional resources for one another   |   |
| Learner | Understanding learning intentions and criteria for success             | <b>5</b> Activating students as the owners of their own learning  |   |

**Fig. 1** Aspects of formative assessment

Crossing the three processes with the different agents (teacher, peer, learner) suggests the framework shown in Fig. 1 (from Wiliam and Thompson 2007), indicating that formative assessment can be conceptualized as consisting of five key strategies:

1. Clarifying and sharing learning intentions and criteria for success<sup>1</sup>;
2. Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding;
3. Providing feedback that moves learners forward;
4. Activating students as instructional resources for one another; and
5. Activating students as the owners of their own learning.

The five types of activity that we had identified earlier as our starting point for our work on formative assessment can therefore be seen as means of enacting these key five strategies. Classroom questioning is merely one way of implementing the second, and comment-only marking is a particular way that teachers might achieve the third. Similarly, peer- and self-assessment are activities that might be used to pursue the fourth and fifth respectively. These are particularly relevant to the development of students' own capacity to learn how to learn and to learner autonomy (Black et al. 2006).

The last of the original types of activity—formative use of summative tests—is more complex. Summative tests (or more properly, tests designed primarily to serve a summative function) provide ways of eliciting evidence of student achievement, and used appropriately, can prompt feedback that moves learning forward. These can also communicate to learners what is and is not valued in a particular discipline, thus communicating criteria for success. Where this has been done, it opens up the possibility of students helping one another, and using the tests as a guide to planning their own revision (Black et al. 2003, Chapter 4).

Wiliam (2007a) gives an extended account of this formulation. It serves to make much clearer the links between the five formative assessment strategies identified in

<sup>1</sup> The importance of “success criteria” and “learning intentions” was emphasised by Clarke (2001).

the framework. The first three will be subject of detailed discussion later in this paper. For the fourth, activating students as instructional resources for one another links to work on collaborative learning (Slavin et al. 2003) and reciprocal teaching (Brown and Campione 1996). Finally activating students as owners of their own learning brings in metacognition (Hacker et al. 1998), motivation (Ryan and Deci 2000), interest (Hidi and Harackiewicz 2000), and attribution (Dweck 2000) as well as self-assessment. Such expansions represent ways in which the understanding of these activities has evolved as they have been implemented, and evaluated, more widely. One outcome of such evolution has been to move from a mere list of activities, with the dangers that this carries of superficial adoption, to a linking of these within a coherent rationale.

In our first formulation of a theory of formative assessment (Black and Wiliam 2006), we focused more directly on issues that had been salient in our earlier work (Black et al. 2003; Wiliam et al. 2004). For the purpose of analysing the processes of change involved in such development, we adopted a framework derived from cultural historical activity theory. That framework helped in the identification of four main themes for discussion, namely:

- Teachers, learners, and the subject discipline;
- The teacher's role and the regulation of learning;
- Feedback and the student-teacher interaction discussed in terms of levels of feedback, the fine-grain of feedback and differentiation;
- The student's role in learning.

In what follows we shall be dealing in more detail with the second and third of these. We shall also refer to some aspects of the other two, but a fuller treatment of these is beyond our present scope.

### 3 Formative assessment, communication and contingency

The purpose of this section is to present a general overview of formative interactions which will lay the basis for meeting all of our five aims rather than serve any one of them directly.

We start by re-stating our definition of formative assessment, drawing both on our earlier definitions (Black and Wiliam 1998b) and the definition of the Assessment Reform Group (ARG 2002), in terms slightly different from, but consistent with both:

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited.

Several features of this definition require elucidation.

The first is the use of the term "instruction". In much of the English-speaking world, this term has a connotation of training, or of didactic approaches to teaching. However, in American English, the term "instruction" means the combination of teaching and learning, and it is this latter sense that we intend here. In this definition

“instruction” refers to any activity that is intended to create learning<sup>2</sup>. The alternative term pedagogy also has drawbacks in suggesting a more narrow approach to the practice of teaching (e.g. it would seem odd to refer to students as ‘pedagogical resources for each other’).

The second is the focus on decisions. We could locate the focus of the definition on the *intentions*, of those involved in instruction, in collecting the evidence, but that would mean that a situation in which evidence was collected, but not used would be formative, which would be unfortunate. Another possibility would be to focus on the resulting *action*: to require that the evidence be used to make instructional adjustments that actually improved learning beyond what would have happened without those adjustments. This, however, is too stringent: given the unpredictability of learning, actions that would appear to be the most likely to produce learning might not do so in a particular situation. An additional problem with such an approach is that it would, in effect, be impossible in practice to establish whether a particular assessment was indeed formative, since it would involve establishing a counter-factual claim: that what actually happened was different from (and better than) what would otherwise have happened (but did not do so). The probabilistic formulation (that the decisions are *likely* to be better) also reflects the fact that even the best designed interventions will not *always* result in better learning for *all* students.

The third feature of the definition concerns the agent of the assessment. While it is clear that in many cases, the decisions will be made by the teacher, the definition also includes peers, or the individual learner, as agents in making such decisions.

The fourth feature is the requirement that decisions are *either* better *or* better founded, than decisions made without the evidence elicited as part of the formative process. The second possibility is included because the formative assessment might, for example, indicate to the teacher that the best course of action is that which the teacher had intended prior to the elicitation of evidence. In this case, formative assessment would not change the course of action, but it would mean that it was better grounded in evidence.

From the definition, it is clear that formative assessment is concerned with the creation of, and capitalization upon, ‘moments of contingency’ in instruction for the purpose of the regulation of learning processes. This might seem to be a very narrow focus, but it helps to distinguish a theory of formative assessment from an overall theory of teaching and learning. However, whilst this focus is narrow, its impact is broad, since how teachers, learners, and their peers create and capitalize on these moments of contingency entails considerations of instructional design, curriculum, pedagogy, psychology and epistemology.

These moments of contingency can be synchronous or asynchronous. Examples of synchronous moments include teachers’ ‘real-time’ adjustments during one-on-one teaching or whole class discussion. Asynchronous examples include teachers’ feedback through grading practices, and the use of evidence derived from homework, or from students’ own summaries made at the end of a lesson (e.g.

---

<sup>2</sup> For those who believe that this merely moves the burden of definition onto the word “learning” we would define learning as an increase, brought about by experience, in the capacities of an organism to react in valued ways in response to stimuli.

‘exit passes’), to plan a subsequent lesson. They might also include responses to work from the students from whom the data were collected, or from other students, or insights learned from the previous lesson or from a previous year.

The responses of teachers can be one-to-one or group-based; responses to a student’s written work is usually one-on-one, but in classroom discussions, the feedback will be in relation to the needs of the subject-classroom as a whole, and may be an immediate intervention in the flow of classroom discussion, or a decision about how to begin the next lesson.

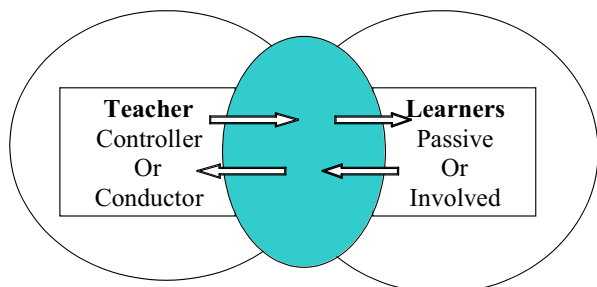
A formative interaction is one in which an interactive situation influences cognition, i.e., it is an interaction between external stimulus and feedback, and internal production by the individual learner. This involves looking at the three aspects, the external, the internal and their interactions. Figure 2 below serves to illustrate the sequence of the argument. The teacher addresses to the learner a task, perhaps in the form of a question, the learner responds to this, and the teacher then composes a further intervention, in the light of that response. This basic structure has been described as initiation–response–evaluation or I-R-E (Mehan 1979), but this structure could represent either a genuinely dialogical process, or one in which students are relegated to a supporting role.

Frequently, the teacher’s use of the I-R-E format involves the teacher asking students to supply missing words or phrases in the teacher’s exposition of the material—a form of extended ‘cloze’ procedure. During such interaction, the teacher’s attention is focused on the correctness of the student’s response—what Davis (1997) terms “evaluative listening”, and subsequent teacher ‘moves’ are aimed at getting the student to make a correct response, through such encouraging responses as, “Almost” or “Nearly”. There is ample evidence that this form of interaction is the norm in most classrooms (Applebee et al. 2003; Hardman et al. 2003; Smith et al. 2004).

The model is meant to apply to more than one-on-one tutoring (which Bloom 1984, regarded as the most effective model of instruction): the shaded area in the centre stands for the classroom where many learners are involved, through hearing the exchange, perhaps by joining in, so there would be many arrows in all directions in this area. This aspect will feature in our later sections.

The process represented in Fig. 2 may be decomposed into several steps: one step is the teacher’s interpretation of the pupils’ responses - this will be discussed in Section 4. The next is to decide on the best response: such decision is first of all a strategic one, in that it can only be taken in the light of the overall purpose for which

**Fig. 2** The three interacting domains of pedagogy



the original task was designed: this will be discussed in Section 5. There is then a tactical decision: how to formulate the detail of the response to best serve the overall strategy: this will be discussed in Section 6.

#### 4 Models for interpretation of the learner's responses

In relation to our five aims, this section will explore connections between formative work and theoretical writing about pupils' learning (our third aim). This will provide a basis for exploring ways of improving practice (our fourth aim).

In a formative mode the teacher's initial prompt is designed to encourage more thought, the learner is more actively involved, and the teacher's work is far less predictable: formative interaction is *contingent*. In such a mode, the teacher's attention is focused on what she or he can learn about the student's thinking from their response—what Davis (1997) terms “interpretive listening”. However, what the learner actually hears and interprets is not necessarily what the teacher intended to convey, and what the teacher hears and interprets is not necessarily what the student intended to convey: the broken arrows in Fig. 2 emphasise this. In a genuinely dialogic process, the teacher's own thinking may come to be modified through the exchange – what Davis terms “hermeneutic listening”.

##### 4.1 The need for a model

Van Lier proposed a model similar to that illustrated in Fig. 2. In discussing the learner's thinking he used the term ‘internal contingency’, but did not pursue his analysis beyond commenting on the rich variety of any learner's mental life and the notorious difficulty of teasing apart the linguistic from the mental, stating that:

From the perspective of contingency it is sufficient to say that speakers, in their use, give evidence of the mental life behind their utterances, by using words such as ‘I believe that..’, ‘it's like x,’ ‘that reminds me of...’ as well as in a myriad of more covert and subtle ways. (van Lier 1996, p. 174)

Perrenoud, in his response to our 1998 review, identified the same problem as follows:

Without a theoretical model of the mediations through which an interactive situation influences cognition, and in particular the learning process, we can observe thousands of situations without being able to draw any conclusions. (p. 95)

Yet insofar as the teacher's focus is on intervention to regulate the learning activity, any intervention has to involve:

...an incursion into the representation and thought processes of the pupil to accelerate a breakthrough in understanding a new point of view or the shaping of a notion which can immediately become operative. (p. 97)

The challenge presented here also becomes evident in other analyses. Sadler (1989) in describing learning as the activity of closing the gap between a learner's present state and the state implied by the learning aim, links this to meta-cognition



and talks of this work taking place “in the act of production” of feedback: that phrase opens up the question of what exactly might be happening in this act of production. In a later discussion (Sadler 1998) that describes the resources that a teacher brings to formative interactions, he points out that:

Teachers bring evaluative skill or expertise in having made judgments about student efforts on similar tasks in the past. [...] In non-convergent learning environments, this automatically exposes teachers to a wide variety of ways in which students approach problem solving, and how they argue, evaluate, create, analyse and synthesise. (p. 81)

There is a formidable problem for teachers, since their feedback needs to be constructed in the light of some insight into the mental life that lies behind the student’s utterances. Indeed, as we have noted elsewhere (Black et al. 2002), it was surprising to us how many teachers had functioned reasonably effectively in the classroom for many years, and were widely regarded as good teachers without having developed such insights:

One of the most surprising things that happened during the early inset sessions was that the participating teachers asked us to run a session on the psychology of learning. In retrospect, perhaps, we should not have been so surprised. We had, after all, stressed that feedback functioned formatively only if the information fed back to the learner was used by the learner in improving performance. But whilst one can work out *after the event* whether or not any feedback has had the desired effect, what the teachers needed was to be able to give their pupils feedback that they knew *in advance* was going to be useful. To do that they needed to build up models of how pupils learn. (p. 14)

The validity of such models will be a necessary, but not sufficient, condition for the effectiveness of the feedback. So our next step is to consider whether any models of the inner mental life can help secure such validity.

#### 4.2 The self-regulated learning model

To meet Perrenoud’s requirement for “theoretical models of the mediation”, we consider theories that come under the general description of self-regulated learning (SRL). A general review of this field (Boekaerts et al. 2005) defines the term as follows:

Self-regulation can be defined as a multi-component, multi-level, iterative self-steering process that targets one’s own cognitions, affects and action, as well as features of the environment for modulation in the service of one’s goals. (p. 150)

In a more detailed review of SRL in education, Boekaerts and Corno (2005) identify three main processes. The first, described as *top-down SR*, has a focus on learning. Here the student pursues the purpose of achieving learning goals that increase resources, i.e. knowledge and both cognitive and social skills. The process is motivated and steered by personal interest, values, and expected satisfaction and

rewards, and may be characterised as the *growth*<sup>3</sup> option. However, it may not be adopted: the student may seek competitive performance goals or prioritise friendship with peers, which a focus on learning goals may put at risk, choosing instead an alternative process, described as *bottom-up SR*. This may be triggered by some types of classroom feedback and reward, or merely by boredom. When cues from the environment have this effect, this second option is adopted—that of giving priority to well-being:

A search for well-being implies that students are more concerned with maintaining or restoring positive feelings than with the pursuit of growth goals. (p. 204)

However, whilst the threats to well-being may lead to such negative effects as aggression, withholding effort, avoidance, or denial, they may also be met with such positive strategies as seeking social support or calling on problem solving strategies previously learnt.

Boekaerts and Corno see SRL as a process of adopting one or other of the *growth track* and the *well-being track*, with the possibility of switching to and fro between the two *en route*. They draw on findings of Vermeer et al. (2001) who:

...also found that students' willingness to maintain learning intentions and persist in the face of difficulty depends on their awareness of and access to volitional strategies (metacognitive knowledge to interpret strategy failure and knowledge of how to buckle down to work). p. 206

Other features relevant to this third element of volitional strategies are helplessness, and failures of emotional control. Thus it is important to help students to acquire positive volitional strategies so that they are not pulled off the *growth track* onto the *well-being track*.

Boekaerts and Corno analyse the relevance of their model to several innovations in classroom work, grouped under cognitive behaviour modification, direct teaching of skills and strategies, and those that take a socio-cultural approach. Here the account draws on some of the literature mentioned in Section 1 and which is discussed further in Sections 5 and 6. Their account does not present an explicit model, i.e., it does not offer any direct help with the task of interpreting a response, in terms that could provide guidance for the teacher on how appropriate feedback may be formulated. However, in discussing top-down SR, they do refer to the model of Winne and Hadwin (1998) stating that it:

...specified the recursively applied forms of metacognitive monitoring and feedback that change information over time (thus influencing goals) as self-regulated learners engage in an assignment. (p. 203)

This model, which owes its origin to information processing theory, was further developed in Greene and Azevedo's (2007) review of more recent SRL studies. Its

<sup>3</sup> Boekaerts and Corno use two terms, *growth* and *mastery*, with both, apparently interchangeable, to stand for a focus on learning achievement; we have preferred the first of these here because *mastery* denotes a sense of completion whereas *growth* does not.

main features, in form simplified here in order bring out features particularly relevant to the present argument, may be represented as follows:

|  |  |
|--|--|
| 1 Identify task                          | A Conditions (of learner and context)        |
| 2 Planning a response                    | B Operations to transform input and own data |
| 3 Enacting a strategy                    | C Standards: criteria for self-appraisal     |
| 4 Adapting: reviewing perhaps re-cycling | D Evaluation                                 |

The model combines two kinds of influences: the left hand column represents the stages 1 to 4 of the production of a response, whilst the right-hand column represents the various resources, A, B, C or D, which might be brought to bear during these stages. Central to the model is the learner's overall control and monitoring function which steers both progress through the stages of production, and the ways in which the resources are deployed—each of 1, 2, 3 and 4 could be associated with any or all of A, B, C, D. There is no implication that these various elements will necessarily be conscious, and all are internal.

Feature A, the *Conditions*, is a broad category comprising the resources available to a person and the constraints inherent in a task or environment. Included are *first* of all cognitive conditions—past experiences recalled, beliefs, dispositions, motivation, and knowledge of the domain, of the current task and of relevant tactics and strategies. *Secondly*, there are task conditions, including resources, time, instructional cues and local context—all external to the person. It follows that *Conditions* influence both the *Operations* and the *Standards* that the learner will deploy.

One *Condition* of particular relevance is the quality of the learners' domain knowledge (and in particular whether it is superficial or deep), which is linked with their personal epistemology. These factors help frame the perception of the task, the precise identification of which is the work done in the first phase of response.

The perspectives formulated at the outset will be transformed in each phase by the *Operations*: these include the searching, assembling, rehearsing and translating processes. Both the ease with which the student can perform these, and the memory capacity on which they draw, are limiting factors here. It is a feature of the model that it is assumed that meta-cognition is not involved in these operations.

Metacognitive faculties are associated with the monitoring and control functions that are exercised through the cognitive *Evaluations*, the latter being formulated in the light of the *Standards*. Many different features may come into play at this stage. The learner's *Standards* will depend in part on their interpretation of the task, on their perception of the criteria and targets for success, on their personal orientation towards the task, and on their view of the time constraints.

For any phase, the overall control and monitoring function may lead to the work being re-cycled, either after the phase itself, or after evaluation of another phase. A learner may revise her/his definition of the task because of a judgment, made in the light of difficulty in phase 3, that it may take too much time, or because the challenge of the original definition implied a high risk of failure, and so start afresh with a revised task definition. On the other hand, if (say) the outcome of phases 1 and 2 is incompetent, but is not perceived as such, then the performance outcome of phase 3 will proceed: if the result then seems to be inadequate, the learner may *Adapt*, i.e., undertake a more comprehensive re-cycling, changing both the *Task*

*Identification* and the choice of *Operation* to enact a very different strategy. Such examples will suffice to indicate that multiple and varied cycles may ensue before an outcome is produced.

Two practical examples may illustrate the issues involved. In the first, a teacher says to a six-year-old drawing a picture of a daffodil: “What is this flower called?” The child answers: “I think it’s called Betty” (Fisher 2005). The child has identified the task in terms of her understanding of the term “called”; her standard for a satisfactory answer is that of everyday talk, not that of a learning discourse in which the distinction between proper and generic names is essential. If the teacher sees the response in this way, feedback may be composed to open up discussion on various meanings of the phrase “is called”. In the second example (Lighthall 1988) a teacher introducing the topic of infinity asks students to say what the term might mean to them. The first to respond is a student, who has such serious behaviour problems that he is having special counselling, saying “I think it’s the back of a *Cream of Wheat* box” to which the teacher responds “Don’t be silly Billy”. In later discussion with the counsellor, Billy explained that he had noticed at breakfast that on the back of the box of the *Cream of Wheat* cereal there was a picture of a man holding this same box, and this picture showed a man holding the same box, and so on. In this example, the student identified the task very accurately, and drew on various resources, including his breakfast experience, to formulate an answer. He was trying to engage because he wanted to overcome past tendencies to misbehave. His criterion for what would count as an adequately clear answer was inadequate, but the real difficulty lay with the teacher’s failure to consider that the answer might be valid and so to ask for more explanation.

Both these examples illustrate the teacher’s problem: on receipt of a response, she has to decide how the student came to make it. Examples of interpretations could be: the student misunderstood a specific feature of the question, or the student has no idea of the sort of answers that are called for, as in the case of the ‘name’ of the flower. The student may have been using some other inadequate *growth* strategy, such as working with a mis-conception, or may have been using a *well-being* strategy, which seems to have been the teacher’s (incorrect) interpretation of Billy’s response. Given an interpretation, the teacher has to act on it, and there are choices here even within any one interpretation e.g. whether to try to point out a perceived flaw in an argument, or to open up an exploratory discussion to develop the reasoning deployed in arguments of this kind, or to work on the student’s resources, e.g. by inviting exploration of various uses of the phrase ‘is called’. Here the accounts of the Winne and Hadwin model fall short—they seem to assume the context of students working on their own and relying only on assessing themselves: the teacher was not mentioned in the original model. Greene and Azvedo partly repair this omission, but the studies that they quote do not pick out work on classroom discourse and do not consider the dynamics of a succession of feedback interventions by the teacher. What the model does underline is that self-assessment is inevitably deployed by learners before any outcome ensues, so the issue is not to initiate self-assessment but to make it more overt as a step to improvement.

It would seem possible to incorporate Boekearts and Corno’s three elements, the processes of growth in learning and of well-being, and the volitional strategies, in this model, for they are *Conditions*, seen as resources, whilst the choice between a

growth and a well-being approach is entailed when proceeding from identifying the task to planning a response. In a response to their review, Winne (2005) expresses the view that all three elements are cognitive strategies, e.g. the learner has to recall past experiences and available resources in deciding which process to adopt, and this is cognitive action, albeit coloured by emotions. An exploration of this argument would go beyond our present purpose: for this purpose, the issue is whether the model can serve as a guide to teachers' contingent actions.

#### 4.3 Practical guidance?

In formulating effective feedback the teacher has to make decisions on numerous occasions, often with little time for reflective analysis before making a commitment. The two steps involved, the diagnostic in interpreting the student contribution in terms of what it reveals about the student's thinking and motivation, and the prognostic in choosing the optimum response: both involve complex decisions, often to be taken with only a few seconds available.

The practical question then is whether the model explored above can give guidance on how formative feedback should be constructed. For example, it can be argued that the detail in the SRL model brings out that an imperfect output may be evidence of a number of different problems. For example, some possibilities are that the learner:

- Has misunderstood the language used, e.g. is baffled by a 'why' question because of failure to understand the meaning of 'why';
- Has misunderstood the whole purpose and context, e.g. assuming that maths is about getting the algorithm right, that all maths problems have one and only one solution, or that history is about telling it "how it really was";
- Has been misled by a superficial aspect of the problem, e.g. thinking that the square root of 0.4 is 0.2;
- Has misunderstood the particular task, e.g. overlooking that the point is to explain, not to give an improved account;
- Is using an inappropriate or ineffective strategy to tackle the task;
- Has not understood the criteria of quality e.g. assuming that a good piece of autobiography should present the correct facts clearly and in order, and has to have accurate punctuation, spelling and paragraphing—thus yielding a dull account devoid of feelings;
- Has given a relevant answer but needs to attempt an explanation at a deeper level, e.g. in a science lesson on plant growth and photosynthesis, where the teacher asked why two plants had grown differently, a student says "this plant in the window grew stronger than that one in the cupboard because it had more light"—whereas the point of the question was to explore explanations of why and how light affects growth (Black et al. 2003, pp. 36–39).

For any of the analysis presented above, we must add a final note of caution, as expressed by von Glasersfeld (1987):

Inevitably, that model will be constructed, not out of the child's conceptual elements, but out of the conceptual elements that are the interviewer's own. It is in this context that the epistemological principle of fit, rather than match is of crucial

importance. Just as cognitive organisms can never compare their conceptual organisations of experience with the structure of an independent objective reality, so the interviewer, experimenter, or teacher can never compare the model he or she has constructed of a child's conceptualisations with what actually goes on in the child's head. In the one case as in the other, the best that can be achieved is a model that remains viable within the range of available experience. (p. 13).

## 5 Feedback strategy and the learning aims

In the practice of instruction the first step is to decide upon the aims and then to plan the activities through which these aims may be realised. The overall purpose of this section is to explore how these decisions differ according to different types of aim and how such differences may change the nature of the formative interactions. This will help with our second aim, of locating our model of formative feedback within more comprehensive theories of pedagogy.

The aims of any instruction are usually a combination of aims specific to the subject and aims directed to improving learning skills. For many teachers, the former are explicit and the latter only implicit. The formative practices, as set out in our discussion of Fig. 1, reflect very general principles of learning, notably social constructivism and meta-cognition. The issues involved are highlighted here first by discussion of two very distinctive programmes of instruction, namely cognitive acceleration and dynamic assessment, in both of which the improvement of learning capacity using an explicit and detailed theory of learning is given priority over subject-specific aims. This is followed by discussion of classroom work in which the subject-specific aims are given more priority. A closing section explores the more general issue of the nature of the teacher's control of learning.

### 5.1 Learning model priority: The cognitive acceleration programmes

The cognitive acceleration (CA) programmes developed by Shayer and Adey (Shayer and Adey 2002; Adey 2005) are a distinctive form of instruction in that they offer a comprehensive innovation programme, with an explicit theoretical basis, a closely specified set of classroom activities, and prescribed pedagogic practices for which teachers require specific training. There is strong evidence that they lead to significant long-term improvements in school achievement, and that the improvements they secure extend beyond the particular subject context in which a programme has been implemented.

The theoretical basis derives from both Piaget and Vygotsky. One central aim is to encourage cognitive growth by creating cognitive conflict, following Vygotsky's dictum from "Mind in Society:

learning which is oriented toward developmental levels that have already been reached is ineffective from the viewpoint of a child's overall development. It does not aim for a new stage of the developmental process, but rather lags behind this process. The only good learning is that which is in advance of development. (Vygotsky 1978, p. 82).

As Chaiklin (2003) makes clear, in reading Vygotsky it is important to understand that Vygotsky drew a clear distinction between learning and development. The latter requires changes in the psychological functions available to the learner, while learning involves the acquisition of new mental capabilities, without changes in the available psychological functions. The zone of proximal development (ZPD) is not, therefore, just a way of describing what a student can do with support, which might be simply learning, it is a description of the *maturing* psychological functions rather than those that already exist. A focus in instruction on the maturing psychological functions is most likely to produce a transition to the next developmental level and “good learning” is that which supports the acquisition of new psychological functions. This careful distinction between learning and development is a central feature of Vygotsky’s work that is often overlooked.

Within the CA framework, metacognition is regarded as a higher level psychological process. By challenging learners to reflect on their own thinking, teachers and their peers help them to make unconscious processes overt and explicit and so making these more available for future use. A key feature of the programmes is that students must learn through dialogue with others, again following Vygotsky’s principle that ideas appear first in the external “social” plane, then become internalised by the individual. These three principles, Cognitive Conflict, Metacognition, and Social Construction, are supplemented by two others.

One of these is a set of schemas of Cognitive Operations, namely seriation, causality, classification, proportionality and probability. The classroom activities are designed to challenge the learner’s capacity to use these. The second is the principle of Bridging, in which learners are encouraged to discuss other contexts in which they might use the same sort of thinking. In the programme for Cognitive Acceleration through Science Education (CASE), teachers are advised to devote one science period every 2 weeks for a period of extending over years 7 and 8 (ages 11 to 13), to specifically designed tasks – these periods are not available for ‘covering’ the normal curricular content.

The emphasis paid to creating cognitive conflict rather than giving answers, to the importance of dialogue to serve the social construction of knowledge, and to metacognition involving learners’ reflection on their own learning, makes it clear that formative assessment practices are an essential feature of these programmes. Indeed, the training process that forms part of the programmes is essential because their adoption requires teachers to engage in such practices, practices which many will find unfamiliar and challenging. Thus, whilst the programme of instruction is distinctive, formative assessment principles lie at the core of its implementation. In SRL terms, the purpose is to change one vital element of the conditions, i.e. the reasoning resources that a learner might bring to any future task.

## 5.2 Learning model priority: Dynamic assessment

Dynamic assessment (DA) is a system with an approach very similar to that of formative assessment (Poehner and Lantolf 2005). The main difference is that it is based on an explicit theory for guiding and interpreting teachers’ work (the Learning Potential Assessment Device of Feuerstein et al. 2003). As for the CA programmes, the model guides the formulation of tasks that are designed to elicit and develop

these two aspects in comprehensive and increasingly sophisticated ways. It also requires that the teacher's responses be guided by the aim of challenging and developing the learner's thinking.

Poehner and Lantolf emphasise that the teacher should not be content with immediate interventions that resolve a particular learning obstacle, but should follow up each success in a sustained and strategic way to build up further the learner's capacity to learn, i.e., exploit to the full the learner's ZPD, pursuing this aim even if this focus is at the expense of the teacher's aims in making progress with the learner's understanding of any particular curriculum topic. They dismiss much of the accounts in the formative literature as too vague because they lack such orientation. However, their judgment that other work on formative assessment lacks a theoretical basis is not accepted by Leung (2007) who has expressed a more optimistic view of the synergy between the two, and points out that Poehner and Lantolf have overlooked accounts of formative assessment which discuss the relevance of the ZPD concept in ways not dissimilar to their own.

In relation to our argument, the DA programme is in the same category as the cognitive acceleration programmes in two respects. It is a comprehensive pedagogy, with a focused purpose not directly related to the conventional curriculum, and it specifies the particular ways in which its specified tasks should be implemented. However, formative assessment practices are central to this implementation.

### 5.3 Learning model priority: Comparisons and contrasts

Other initiatives might be analysed in a similar fashion. An example is the innovation described by White and Frederiksen (1998) that draws on a particular cognitive model (called ThinkerTools), applies it within science lessons, and produces impressive evidence of improvement.

What is common to most of the approaches discussed here is that they all include particularly effective versions of classroom discourse that, by careful choices of the tasks, the cues, and the feedback, exploit the potential fluidity of the learner's strategies to make significant and cumulative changes in them. Indeed, such programmes may be classified as formative assessment exercises which rely upon interactive dialogue, both between teachers and students and between students themselves, distinguished by the very specifically targeted nature of the reasoning tasks that these methods deploy, the articulation of these in a sustained and coherent sequence, and the specification of aims by the light of which any feedback responses may be guided.

Seen from the perspective of the CA and DA programmes, any teacher using formative and interactive dialogue for normal subject teaching, yet using feedback informed by attention to the issues highlighted in the SRL model, is engaged in a more diffuse and subject-specific form of thinking skills programme. The main types of activity summarised in Section 1 above do fit easily within these programmes—indeed, they are essential to their implementation.

### 5.4 Subject content priority: Lesson plans

Whatever their strategy, teachers have to start a lesson with an “opening move”. In many classrooms, this will be an exploratory question, designed to elicit students'



existing conceptions. However, the way in which teachers then proceed may differ profoundly with the broader cultural context within which they work. For example, account must be taken of the complexities introduced by the requirement of the teacher to assume responsibility for organizing the learning of a large number of students (20 to 40 in the developed world, often much larger in the developing world). Of the many possibilities within this broader agenda, we expand on one example as follows.

In many communities all over the world, there is increasing acceptance of a canonical lesson design which may now be sufficiently widespread to qualify as a “signature pedagogy” (Shulman 2005). The lesson begins with a “big question” (*hatsumon* in Japanese) which has been carefully designed to lead students towards the intended outcomes (however broadly they may be defined)<sup>4</sup>. Students are asked to work on this question in pairs or small groups, and then the teacher conducts a whole class session in which different groups present their proposals. Typically, the teacher then conducts a whole-class discussion, which is termed “*neriage*” in Japanese; this word, whose literal meaning is “kneading”<sup>5</sup>, is used in education to describe:

the whole class discussion phase of structured problem solving. It is the core of teaching through problem solving. This happens after students have shared various solution strategies. During this phase, students, carefully guided by the teacher, critically analyze, compare and contrast the shared ideas. They will consider issues like efficiency, generalizability, and similarity to previously learned ideas. (Takahashi 2008 p. 8)

In conducting the *neriage* session, the teacher must balance a range of different concerns, some of which may conflict with the others. The teacher must retain the focus on the learning. If student contributions raise new possibilities, the teacher has to make split-second decisions whether to follow the new thread, or bring the conversation back to where the teacher intended. The pressure to value every contribution is strong, since as well as advancing the learning of the whole class, the teacher will be seeking to minimize the sense of rejection a student might feel if her or his contribution is dismissed. In this respect it is interesting to note that Bromme and Steinbring (1994) discovered in their “expert-novice” analysis of two mathematics teachers that the novice teacher tended to treat students’ questions as being from individual learners, while the expert teacher’s responses tended to be directed more to a “collective student”. In other words, for each contribution, the expert teacher sought to draw out implications for the learning of the whole class, rather than for each individual student.

### 5.5 The teacher’s control of learning

The above studies take for granted several larger issues about the aims and the contexts of classroom work. To consider first the aims, we may assume that at the beginning of any classroom discourse the teacher will have some form of learning intention. We do

<sup>4</sup> Examples of such questions in mathematics may be found in Hodgen and Wiliam (2006) and Swann (2006).

<sup>5</sup> ‘*neriage*’ originally applied to the technique of layering, cutting and re-combining different colours of clay to produce a block with intricate patterns.

not, however, assume that the teacher has anything as narrow as a single pre-determined goal for all students. The teacher may be happy for different students to be working towards different goals, and the teacher may not herself be clear about what the learning outcomes achieved will turn out to be. However, it is our contention that in such situations, the teacher does have learning intentions however implicit, for otherwise the situation would be that “anything goes”. This is an important point, because some authors (e.g., Knight 2008) have suggested that requiring the teacher to have learning intentions, and to place the responsibility for the creation of learning with the teacher, is somehow to undermine the creation of student autonomy. Such an argument, if taken to the extremes characteristic of some forms of “discovery learning”, seems to us to be based on an inappropriate model of the nature of the respective responsibilities of the teacher and the learner. We would accept that in many situations, particularly those in which high-stakes tests play an important role, teachers often behave as if they believed that they could do the learning for the learner, with disastrous consequences. At the other extreme, however, are the practitioners who claim not to teach, but to merely act as other participants, or to chair the learning discussions. This seems to us an abdication of the teacher’s responsibility. While the teacher cannot do the learning for the learner, the teacher can engineer situations in which the opportunities for the learner either to learn, and/or to develop learning autonomy, are maximized. And while the voice of the student should be taken into account in the engineering of such learning environments, it is fatal if the teacher does not engineer into the learning environment the essence of the discipline. The teacher must be accountable to the students in terms of taking on board, as far as reasonably practicable, the students’ needs, preferences, and so on, but they must also be accountable to the discipline into which the students are being enculturated so that they can eventually operate as effective learners in that discipline.

Whether it is chosen at a strategic or tactical level, the task or problem put to the student(s) cannot be taken for granted: if well chosen, it can engage the students and help draw them in to a learning discussion. However, it is hard to formulate any general rules or recipes. Moreover, as any particular task is implemented, the ‘tuning’ of that task to the responses can also be very challenging for teachers—Dillon (1994) presents detailed guidance to teachers about this issue. In some situations, the teacher will be trying to get all students to the same goal, such as understanding the notion of equivalent fractions. In other situations, the teacher’s aims may be better characterized as having a broad horizon of acceptable outcomes: Hodgen and Marshall’s (2005) study brings out these differences in their comparison of mathematics with English classrooms. Different students in the same class might be working towards different goals, and this might be perfectly acceptable to the teacher. However, it will also often be the case that some students appear to be working towards an outcome that is too far from the teacher’s intention, and in such a case, the teacher is likely to try to draw the student towards a trajectory that is more in keeping with the teacher’s intention.

## 6 Feedback tactics

The explorations in Section 5 have inevitably raised issues about the way in which any strategic intention is implemented in the immediacy of school work. Two

aspects of this issue will be discussed in this section, the first about evidence of the effectiveness of different types of feedback, the second about the relationship of feedback decisions to the development of students' participation in learning discourse. Thus the discussion in this section is aligned particularly with our fourth aim, which is to produce ideas that might lead to improved practice

### 6.1 Categories of feedback

A significant contribution to this topic has come from Hattie and Temperley's (2007) review. The first section of their paper reviews the quantitative evidence for the different types of feedback, using meta-analyses to derive average effect sizes. They obtain high effect sizes in cases where students receive information about the task and about how to perform it more effectively, lower effects for interventions which focus on target-setting, and far lower effects where only praise, or rewards, or punishment are given.

The detailed patterns into which they organize these data are then deployed in the authors' second section, in which they present a classification of feedback into four different types. The first type is task-level feedback, which focuses on faults in the interpretation of the task or in the outcome produced. The second is about the main process needed to understand/perform a task: such feedback should be related to the student's own error-detection strategies, and has to serve as an 'advance-organiser', giving cues to lead to better strategies. The third focuses on the self-regulation level—the self-monitoring, directing, and regulating of actions.

Hattie and Temperley note that willingness to engage with such feedback depends on the 'transaction costs' involved, and that feedback is most effective when learners' confidence in their production is high (even if it turns out to be wrong), but may be ignored where confidence is low. Unclear negative feedback where there is uncertain self-image can lead to poor performance especially if it contributes to what they call the learner's "capricious chaotic reinforcement history". Provided they seek instrumental help and do not look either for answers, or to be told 'how to do it', help-seeking interventions by students can have positive benefits'; however, such helpful interventions are not often made (see Swift et al. 1988; Smith et al. 2004).

The authors make use of Butler and Winne's (1995) paper, but do not relate their analysis to Winne and Hadwin's (1998) model. The three types are all related to the internal components represented in the SRL model, but their perspective differs in that they are analyzing the teacher's feedback actions, i.e. the external and inter-subjective aspects of feedback interaction, rather than relating them to an explicit model of the learner's thinking.

Many valuable lessons emerge from the authors' distillation of findings from a wide range of data, and their categories can be used to classify the surface features of different examples of feedback. However, the diagnosis by the teacher or peers which must precede the formulation of any feedback cannot be restricted to one of these types. The point of the SRL model is to bring out the various elements of self-regulation, and the many ways in which they might interact, which between them lead to the outcome, so that the most appropriate feedback

might be formulated. For example, feedback on understanding of the task may have to be linked with feedback on the learner's understanding of the criteria used in his/her own self-regulation, or on the choice of strategy made in the light of that understanding, and so on.

It does not follow that learners will work on any feedback; both the commitment and the confidence of the student will affect the response, and in this respect the following category, the fourth identified by Hattie and Temperley, is relevant. This concerns issues of personal evaluations and affect, including such feedback features as praise and judgment. There is not, from the evidence, a clear indication for choosing between positive and negative feedback: the former can enhance motivation, whilst the latter may, depending on the recipient's self-belief, be accepted as a challenge which helps by triggering self-regulation.

What is not discussed in Hattie and Temperley's review is the effects of feedback on the learning orientation of the learner. The research studies of Ruth Butler (1987, 1988) have shown that giving marks or grades, or otherwise focusing on judgment or competition, as part of feedback can inhibit the learner's attention to any substantive advice on improvement. Moreover, such feedback can actually have a damaging effect on the learning orientation of the learner, promoting ego- or performance-oriented orientations as opposed to task-orientation. Arguments of a similar type are developed in detail by Dweck (2000), and these reflect the distinction discussed in Section 4 above between the *growth* and the *well-being* processes, bringing out the way in which different types of feedback can influence the volitional strategy and the switching from one process to the other.

Overall, the Hattie and Temperley study does not suggest that empirical evidence indicates any need for a revision of the Winne and Hadwin model. Most of the empirical lessons can be seen as consistent with teacher feedback formulated in the light of that model, whilst providing additional guidance, derived from that evidence, about ways to optimize that feedback. However, neither of these two studies pays serious attention to the external world of classroom discourse and indeed much of what they present can be interpreted as dealing with transactions in which a teacher is interacting with the individual student about a piece of written work. It requires some imagination to connect the lessons with the dynamic interactive environment of a teacher working with a class of 30 students. Furthermore, they do not consider directly reverse feedback, i.e. the feedback of student to teacher.

## 6.2 Promoting discourse

Van Lier (1996) maps out a spectrum of types of class discussion, from the monologue at one extreme to the everyday free-flowing social chat at the other: a rather similar scheme, but more restricted in range, has been proposed by Alexander (2006). It is clear that formative assessment cannot flourish at either end of this spectrum. The optimum balance for different school subjects may lie at different points along it, between relaxed freedom and strongly steered dialogue.

A detailed analysis of classroom dialogue may be undertaken from a variety of perspectives. In the multi-disciplinary study reported by Dillon (1988), extensive transcripts of five classroom discussions were the subject of 12 different types of

analysis and commentary: six contributions analysed the samples in the light of such different disciplines as philosophy, cognitive psychology and socio-linguistics, whilst six others appraised the pieces from different pedagogical perspectives, under such headings as *models of discussion*, *questions and wait-time*, and *questioning vs. student initiatives*; the last of these, by Wood and Wood, illustrates a type of detailed analysis of teachers moves which would be most useful in the further development of the skills needed for formative dialogue.

One feature of Dillon's examples is that the pupils' talk takes up at least as much of the time as the teacher's talk, and in these episodes the pupils' talk in paragraphs. By contrast, two short samples published as part of the formative assessment work reported in Black et al. (2003, ch. 4) show one episode in which pupils' contributions amount to no more than two- or three-word phrases, whilst in a later, and more formative, example from the same teacher, pupils' contributions take the form of complete sentences, using terms such as 'think' and 'because' which indicate reasoned dialogue. In the first of these the teacher is merely seeking the expected right answer, in the second he acts more to steer the conversation by first collecting some contributions, summarising them and asking for more comments on the differences between these, leaving his own guiding judgments and challenges until many have had time to contribute. Nevertheless, it is clear that the conversation is being steered in a definite direction, a feature which is far less clear in Dillon's examples. However, some of the latter are discussions of such topics as family relationships or the school's ban on smoking, whereas the former is about the meaning of a scientific concept.

Further development of such studies could help to develop a more fine-grained understanding of classroom discourse, and might be of value in giving detailed guidance to teachers of the type formulated by Dillon (1994). However, none of them is directly related to models of formative assessment similar to the one that we set out in Fig. 1, which identifies, in its five-strategy model, the key processes of formative assessment. As will be clear from our discussion of Fig. 1, none of these occurs in isolation. While particular instructional episodes may highlight one of the strategies, others are inevitably involved. For example, where students are active as owners of their own learning, this will require them to have some idea about what they are trying to achieve, as will students who are active as instructional resources for one another. This interplay between strategies is particularly complex in the conduct of classroom discourse. What should be emphasised here is that the task of using learners' responses to catalyse their further involvement in a learning discourse between peers is a far more complex task for the teacher than dealing with a single student's response in an individual tutorial.

We shall not consider here the practices of feedback for written work or in peer-assessment. Feedback comments on written work can set up a conversation, a form of dialogue which presents problems similar to those of whole-class dialogue but simpler partly because the teacher has time to consider how best to respond, and partly because subsequent interaction, in writing or in discussion, will be a one-to-one interchange. For peer-assessment, the teacher in classroom interaction can model for learners the way they should interact with one another, but far more will be involved (Blatchford et al. 2006; Mercer et al. 2004).

## 7 Conclusions

What has been presented here as a theory attempts to fill a conceptual gap in the literature on formative assessment. It does so only as a first step, establishing that any theory must bring into relationship the three spheres, the teacher's agenda, the internal world of each student, and the inter-subjective; these between them map the territory. In doing this, we have tried to bring together literature from diverse traditions, each seeming to have its own distinctive aims and limited sphere of concern. We have judged that the task of developing and exploiting the various linkages between them will be profitable in the formulation of a theoretical framework that may serve to illuminate the complexities of educational practice.

Our approach indicates that any evidence of formative interaction must be analysed as reflecting a teacher's chosen plan to develop learning, the formative interactions which that teacher carries out contingently within the framework of that plan—as realised in the social world of the classroom and school—and the internal cognitive and affective models of each student of which the responses and broader participation of students provide only indirect evidence.

The framework sketched out here serves the *first of the aims* set out in Section 1, in that it gives a conceptual re-alignment of the practices currently assembled under the umbrella of formative assessment. However, the focus here has been on the first three numbered items in Fig. 1, with some attention to the other two. A more detailed discussion of self-assessment and peer-assessment is beyond the scope of this paper.

Our *second aim* was to locate formative interactions within more comprehensive theories of pedagogy. Perrenoud's concept of regulation is crucial here. He explained this concept in terms of two levels:

I would like to suggest several ways forward, based on distinguishing two levels of the management of situations which favour the interactive regulation of learning processes:

- the first relates to the setting up of such situations through much larger mechanisms and classroom management.
- the second relates to interactive regulation which takes place through didactic situations. (p. 92)

The first of these is a key feature of pedagogy, but only the second involves formative interaction. Overall, the argument is that the collection of formative practices, whether or not informed by models such as the SRL model, or the cognitive acceleration programmes, or dynamic assessment approaches, are not alternatives to a formative approach—they all include it. Where they differ is in Perrenoud's first level of management, i.e. in the design of their learning tasks, and in the over-riding learning aims which serve to determine the priorities towards which the learning interactions are directed (as is made clear on our Section 5). That is to say, they differ in the model of pedagogy that is the basis for their design and within which formative assessment is, to a greater or lesser extent, implemented. Thus they all involve the teacher in the difficult task of feedback for learning, but in different ways. This aspect is clarified by Threfall's (2005) suggestion that teachers

may need to exercise ‘contingent assessment planning’, i.e., limiting the teaching plan so that the interactions are kept within the range of alternatives that they can foresee and be prepared for (see also Ciofalo and Wylie 2006; Wylie and William 2006, 2007). For example, the focused nature of (say) the Cognitive Acceleration programmes, together with their focussed professional development, makes such contingent planning simpler because the range of appropriate responses is correspondingly narrowed and the training can anticipate the main characteristics of students’ reactions.

Our *third aim* was to link and critique other theoretical writing. What the above exploration has opened up for us is the ways in which the self-regulated learning literature pays scant attention to learning in the context of discourse, whether with a teacher or a member of a group, whilst the literature on discourse does not try to incorporate and link to literature on the individual’s learning. Socio-cultural theorisation of classrooms is clearly relevant to consideration of classroom discourse but the approach we illustrate in Fig. 1 also recognises the need to consider the learner as an individual thinker.

Our *fourth aim* was to produce ideas that might lead to improved practice. There does seem to be potential here, e.g. in more sophisticated guidance for teachers to help them both to interpret students’ contributions, and to match their contingent responses to the priority of purpose which they intend. Such work might well start from existing analyses of classroom discourse mentioned in Section 6. It might also help to make clear the distinction between teachers’ broad approach to instruction, including the learning theories to which they subscribe, and the specific issues involved in implementing a formative approach. It could illustrate and emphasise that such an approach cannot prosper if the up-stream planning does not provide a context favourable to it.

At a finer grained level, and strongly contingent, is the prospect of guiding the individual choices of feedback that a teacher has to make on numerous occasions, often with little time for reflective analysis before making any commitment. The two steps involved, the diagnostic, in interpreting the student contribution in terms of what it reveals about the students thinking and motivation, and the prognostic, in choosing the optimum response, both involve complex decisions, often to be taken with only a few seconds available. It might be attractive to attempt to compose a list of rules from an extensive collection of examples like the few presented in Section 6. However, there would be dangers. The first is that a response cannot be interpreted outside the context within which it has been produced. Another is that a rule that might apply in one subject, or in a lesson that was pursuing a particular learning aim within that subject, might be quite inappropriate for another subject or for an interaction that was being orchestrated with a different aim in mind.

We do not claim that the focus on formative assessment is the only, or even the most appropriate, way of looking at interactions in whole-class discussions. Boaler and Humphreys (2005) have highlighted the importance of social norms in the creation of effective classroom environments, and the work of Paul Cobb and his colleagues have underlined the importance of subject-specific norms (e.g., McClain and Cobb 2001). We would certainly accept that what counts as a good explanation in the mathematics classroom would be different from what counts as a good explanation in the history classroom, although they would also share certain commonalities.

We do claim, however, that a focus on the creation of, and capitalization upon, moments of contingency in whole-class discussion, and on the five key strategies defined in Section 1 above, provide lenses that are useful for both teachers and researchers. In particular, for practitioners the fact that moments of contingency create the possibilities for whole class discussion to be improved provide a point of leverage that seems to us uniquely powerful. Rather than ignoring issues of psychology, curriculum and pedagogy, such a focus allows the teacher to engage with these issues in a way that is directly and immediately relevant to their practice.

Thus it is clear that the complexity of the situations in which formative feedback is exchanged is such that they can only be understood in terms of the several theoretical perspectives required to explore the different types of issue involved. These might variously illuminate the formative aspects involved, or, more likely, the broader theory of pedagogy within which the formative dimension is located. There is ample room to develop such considerations, i.e. to pursue our *fourth aim* of developing further lines of enquiry. Whilst this will not be further explored here, we draw attention to such issues as the difference between the different epistemologies and cultures of the various school subjects (touched on in Section 5), and to the differences between the learning needs of (say) pre-school children and undergraduate specialists.

## References

- Adey, P. (2005). Issues arising from the long-term evaluation of cognitive acceleration programmes. *Research in Science Education*, 35, 3–22.
- Alexander, R. (2006). *Towards dialogic thinking: Rethinking classroom talk*. York: Dialogos.
- Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion based approaches to developing understanding: classroom instruction and student performance in middle and high school English. *American Educational Research Journal*, 40(3), 685–730.
- ARG (2002). *Assessment for Learning: 10 Principles*. Available on the Assessment Reform Group website: [www.assessment-reform-group.org.uk](http://www.assessment-reform-group.org.uk).
- Black, P. (2007). Full marks for feedback. *Make the Grade: Journal of the Institute of Educational Assessors*, 2(1), 18–21 (Spring).
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, 5(1), 7–73.
- Black, P., & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. London: GL Assessment.
- Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81–100). London: Sage.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). *Working inside the black box: Assessment for learning in the classroom*. London: GL Assessment.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.
- Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: a theoretical inquiry. *Research Papers in Education*, 21(2), 119–132.
- Blatchford, P., Baines, E., Rubie-Davies, C., Bassett, P., & Chowne, A. (2006). The effect of a new approach to group-work on pupil-pupil and teacher-pupil interaction. *Journal of Educational Psychology*, 98, 750–765.
- Bloom, B. S. (1984). The search for methods of instruction as effective as one-to-one tutoring. *Educational Leadership*, 41(8), 4–17.
- Boaler, J., & Humphreys, C. (2005). *Connecting mathematical ideas: Middle school video cases to support teaching and learning*. Portsmouth, NH: Heinemann.



- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology, 54*(2), 199–231.
- Boekaerts, M., Maes, S., & Karoly, P. (2005). Self-regulation across domains of applied psychology: Is there an emerging consensus? *Applied Psychology, 54*(2), 149–154.
- Bromme, R., & Steinbring, H. (1994). Interactive development of subject matter in the mathematics classroom. *Educational Studies in Mathematics, 27*(3), 217–248.
- Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: on procedures, principles, and systems. In L. Schauble, & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 291–292). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest and performance. *Journal of Educational Psychology, 79*(4), 474–482.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology, 58*(1), 1–14.
- Butler, D., & Winne, P. H. (1995). Feedback as self-regulated learning: a theoretical synthesis. *Review of Educational Research, 65*, 245–281.
- Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. In A. Kozulin, B. Gindis, V. S. Ageyev, & S. M. Miller (Eds.), *Vygotsky's educational theory and practice in cultural context* (pp. 39–64). Cambridge: Cambridge University Press.
- Ciofalo, J., & Wylie, C. E. (2006). Using diagnostic classroom assessment: one question at a time. *Teachers College Record*, January 10, 2006, <http://www.tcrecord.org>, ID Number: 12285.
- Clarke, S. (2001). *Unlocking formative assessment*. London: Hodder and Stoughton.
- Davis, B. (1997). Listening for differences: an evolving conception of mathematics teaching. *Journal for Research in Mathematics Education, 28*(3), 355–376.
- Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.
- Dillon, J. T. (1988). *Questioning and discussion: A multi-disciplinary study*. New York: Ablex.
- Dillon, J. T. (1994). *Using discussion in classrooms*. London: Open University Press.
- Feuerstein, R., Falik, L., Rand, Y., & Feuerstein, R. S. (2003). *Dynamic assessment of cognitive modifiability*. Jerusalem: ICELP Press.
- Fisher, R. (2005). *Teaching children to learn* (2nd ed.). London: Continuum.
- Greene, J. A., & Azvedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research, 77*(3), 354–372.
- Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.) (1998). *Metacognition in educational theory and practice*. Mahwah, NJ: Erlbaum.
- Hardman, F., Smith, F., & Wall, K. (2003). "Interactive whole class teaching" in the national literacy strategy. *Cambridge Journal of Education, 33*(2), 197–215.
- Hattie, J., & Temperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.
- Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70*(2), 151–179.
- Hodgen, J., & Marshall, B. (2005). Assessment for learning in mathematics and English: Contrasts and resemblances. *The Curriculum Journal, 16*(2), 153–176.
- Hodgen, J., & Wiliam, D. (2006). *Mathematics inside the black box*. London: GL Assessment.
- James, M., Black, P., Carmichael, P., Drummond, M. -J., Fox, A., MacBeath, J., et al. (2007). *Improving learning how to learn in classrooms, schools and networks*. London: Routledge.
- Knight, O. (2008). Create something interesting to show that you have learned something. *Teaching History, 131*, 17–24.
- Leung, C. (2007). Dynamic assessment: assessment for or as teaching? *Language Assessment Quarterly, 4* (3), 257–278.
- Lighthall, F. F. (1988). An organization watcher's view of questioning & discussion. In J. T. Dillon (Ed.), *Questioning and discussion: A multidisciplinary study* (pp. 135–153). New York, NY: Ablex.
- McClain, K., & Cobb, P. (2001). An analysis of development of sociomathematical norms in one first-grade classroom. *Journal for Research in Mathematics Education, 32*(3), 236–266.
- Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.
- Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal, 30*(3), 359–377.

- Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. *Assessment in Education: Principles, Policy and Practice*, 5(1), 85–102.
- Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9(3), 233–265.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28, 4–13.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54–67.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Sadler, D. R. (1998). Formative assessment: revisiting the territory. *Assessment in Education*, 5(1), 77–84.
- Shayer, M., & Adey, P. (Eds.) (2002). *Learning intelligence: cognitive acceleration across the curriculum 5 to 15 years*. Milton Keynes: Open University Press.
- Shulman, L. S. (2005). *The signature pedagogies of the professions of law, medicine, engineering, and the clergy: potential lessons for the education of teachers*. Paper presented at the National Science Foundation Mathematics and Science Partnerships Workshop: Teacher Education for Effective Teaching and Learning held at National Research Council Center for Education, Irvine, CA.
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. M. (2003). Cooperative learning and achievement. In W. M. Reynolds & G. J. Miller (Eds.), *Handbook of psychology, vol 7: Educational psychology* (pp. 177–198). Hoboken, NJ: Wiley.
- Smith, F., Hardman, F., Wall, K., & Mroz, M. (2004). Interactive whole class teaching in the National Literacy and Numeracy strategies. *British Educational Research Journal*, 30(3), 395–411.
- Swan, M. (2006). *Collaborative learning in mathematics: A challenge to our beliefs and practices*. London: National Institute for Advanced and Continuing Education (NIACE) for the National Research and Development Centre for Adult Literacy and Numeracy (NRDC).
- Swift, J. N., Gooding, C. T., & Swift, P. R. (1988). Questions and wait time. In J. T. Dillon (Ed.), *Questioning and discussion: A multidisciplinary study* (pp. 192–211). New York, NY: Ablex.
- Takahashi, A. (2008). *Neriage: An essential piece of a problem-based lesson. Teaching through problem solving: A Japanese approach*. Paper presented at the Annual conference of the National Council of Teachers of Mathematics. Salt Lake City, UT.
- Threlfall, J. (2005). The formative use of assessment information in planning—The notion of contingent planning. *British Journal of Educational Studies*, 53(1), 54–65.
- van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. Harlow: Pearson Education.
- Vermeer, H., Boekaerts, M., & Seegers, G. (2001). Motivational and gender differences: Sixth-grade students' mathematical problem-solving behaviour. *Journal of Educational Psychology*, 92(2), 308–315.
- von Glasersfeld, E. (1987). Learning as a constructive activity. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 3–17). Hillsdale, NJ: Erlbaum.
- Vygotsky, L. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.
- White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modelling, and metacognition. Making science accessible to all students. *Cognition and Instruction*, 16(1), 3–118.
- Wiliam, D. (2000). Formative assessment in mathematics part 3: the learner's role. *Equals: Mathematics and Special Educational Needs*, 6(1), 19–22.
- Wiliam, D. (2007a). Keeping learning on track: classroom assessment and the regulation of learning. In F. K. Lester Jr. (Ed.), *Second handbook of mathematics teaching and learning* (pp. 1053–1098). Greenwich, CT: Information Age.
- Wiliam, D. (2007b). Content then process: Teacher learning communities in the service of formative assessment. In D. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp. 182–204). Bloomington, IN: Solution Tree.
- Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Mahwah, NJ: Erlbaum.
- Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education: Principles, Policy and Practice*, 11(1), 49–65.
- Winne, P. H. (2005). Key issues in modelling and applying research on self-regulated learning. *Applied Psychology*, 54(2), 232–238.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Erlbaum.

- 
- Wylie, E. C., & Wiliam, D. (2006). *Diagnostic questions: is there value in just one?* Paper presented at the Annual Meeting of the National Council on Measurement in Education. San Francisco, CA.
- Wylie, E. C., & Wiliam, D. (2007). *Analyzing diagnostic questions: what makes a student response interpretable?* Paper presented at the Annual Meeting of the National Council on Measurement in Education. Chicago, IL.