Formative Assessment in Mathematics

Part 2: Feedback

Dylan Wiliam

Equals: Mathematics and Special Educational Needs **5**(3) 8-11; Autumn 1999

Introduction

This is the second of three articles reviewing the effectiveness of formative assessment, summarising the findings of a review of over 200 studies into formative assessment (Black and Wiliam, 1998). The focus of this article is the effect that feedback to learners about their past performance has on their future performance and their attitudes towards their learning. In particular, what kinds of feedback are the most effective in promoting learning?

The quality of feedback

Ruth Butler (1998) investigated the effectiveness of different kinds of feedback on 132 year 7 students in 12 classes in 4 Israeli schools. For the first lesson, the students in each class were given a booklet containing a range of divergent thinking tasks. At the end of the lesson, their work was collected in. This work was then marked by independent markers. At the beginning of the next lesson, two days later, the students were given feedback on the work they had done in the first lesson. In four of the classes students were given marks (which were scaled so as to range from 40 to 99) while in another four of the classes, students were given comments, such as "You thought of quite a few interesting ideas; maybe you could think of more ideas". In the other four classes, the students were given both marks and comments.

Then, the students were asked to attempt some similar tasks, and told that they would get the same sort of feedback as they had received for the first lesson's work. Again, the work was collected in and marked.

Those given only marks made no gain from the first lesson to the second. Those who had received high marks in the tests were interested in the work, but those who had received low marks were not. The students given only comments scored, on average, 30% more on the work done in the second lesson than on the first, and the interest of all the students in the work was high. However, those given both marks and comments made *no gain* from the first lesson to the second, and those who had received high marks showed high interest while those who received low marks did not.

In other words, far from producing the best effects of both kinds of feedback, giving marks alongside the comments completely washed out the beneficial effects of the comments. The use of both marks and comments is probably the most widespread form of feedback used in the United Kingdom, and yet this study (and others like it—see below) show that it is no more effective than marks alone. In other words, if you are going to grade or mark a piece of work, you are wasting your time writing careful diagnostic comments.

A clear indication of the role that ego plays in learning is given by another study by Ruth Butler (1987). In this study, 200 year 6 and 7 students spent a lesson working on a variety of divergent thinking tasks. Again, the work was collected in and the students were given one of four kinds of feedback on this work at the beginning of the second lesson (again two days later):

a quarter of the students were given comments;
a quarter were given grades;
a quarter were given praise; and
a quarter were given no feedback at all.

The quality of the work done in the second lesson was compared to that done in the first. The quality of work of those given comments had improved substantially compared to the first

lesson, but those given grades and praise had made no more progress than those given absolutely no feedback throughout their learning of this topic.

At the end of the second lesson, the students were given a questionnaire about what factors influenced their work. In particular the questionnaire sought to establish whether the students attributed successes and failures to themselves (called ego-involvement) or to the work they were doing (task-involvement). Examples of ego- and task-involving attributions are shown in table 1.

Attribution of	Ego	Task
Effort	To do better than others To avoid doing worse than others	Interest To improve performance
Success	Ability Performance of others	Interest Effort Experience of previous learning

Table 1: ego- and task-related attributions

Those students given comments during their work on the topic had high levels of task-involvement, but their levels of ego-involvement were the same as those given no feedback. However, those given praise and those given grades had comparable levels of task-involvement to the control group, but their levels of ego-involvement were substantially higher. The only effect of the grades and the praise, therefore, was to increase the sense of ego-involvement without increasing achievement.

This should not surprise us. In pastoral work, we have known for many years that one should criticize the behaviour, not the child, thus focusing on task-involving rather than ego-involving feedback. These findings are also consistent with the research on praise carried out in the 1970s which showed clearly that praise was not necessarily 'a good thing'—in fact the best teachers appear to praise slightly less than average (Good and Grouws, 1975). It is the quality, rather than the quantity of praise that is important and in particular, teacher praise is far more effective if it is infrequent, credible, contingent, specific, and genuine (Brophy, 1981). It is also essential that praise is related to factors within an individual's control, so that praising a gifted student just for being gifted is likely to lead to negative consequences in the long term.

The timing of feedback is also crucial. If it is given too early, before students have had a chance to work on a problem, then they will learn less. Most of this research has been done in the United States, where it goes under the name of 'peekability research', because the important question is whether students are able to 'peek' at the answers before they have tried to answer the question. However, a British study, undertaken by Simmonds and Cope (1993) found similar results. Pairs of students aged between 9 and 11 worked on angle and rotation problems. Some of these worked on the problems using Logo and some worked on the problems using pencil and paper. The students working in Logo were able to use a 'trial and improvement' strategy which enabled them to get a solution with little mental effort. However, for those working with pencil and paper, working out the effect of a single rotation was much more time consuming, and thus the students had an incentive to think carefully, and this greater 'mindfulness' led to more learning.

The effects of feedback highlighted above might suggest that the more feedback, the better, but this is not necessarily the case. Day and Cordon (1993) looked at the learning of a group of 64 year 4 students on reasoning tasks. Half of the students were given a 'scaffolded' response when they got stuck—in other words they were given only as much help as they needed to make progress, while the other half were given a complete solution as soon as they got stuck, and then given a new problem to work on. Those given the 'scaffolded' response learnt more, and retained their learning longer than those given full solutions.

In a sense, this is hardly surprising, since those given the complete solutions had the opportunity for learning taken away from them. As well as saving time, therefore, developing skills of 'minimal intervention' promote better learning.

Sometimes, the help need not even be related to the subject matter. Often, when a student is given a new task, the student asks for help immediately. When the teacher asks, "What can't you do?" it is common to hear the reply, "I can't do any of it". In such circumstances, the student's reaction may be caused by anxiety about the unfamiliar nature of the task, and it is frequently possible to support the student by saying something like "Copy out that table, and I'll be back in five minutes to help you fill it in". This is often all the support the student needs. Copying out the table forces the student to look in detail at how the table is laid out, and this 'busy work' can provide time for the student to make sense of the task herself.

The consistency of these messages from research on the effects of feedback extends well beyond school and other educational settings. A review of 131 well-designed studies in educational and workplace settings found that, on average, feedback did improve performance, but this average effect disguised substantial differences between studies. Perhaps most surpisingly, in 40% of the studies, giving feedback had a negative impact on peformance. In other words, in two out of every five studies, giving people feedback on their performance made their performance worse than if they were given no feedback on their performance at all! On further investigation, the researchers found that feedback makes performance worse when it is focused on the self-esteem or self-image (as is the case with grades and praise). The use of praise can increase motivation, but then it becomes necessary to use praise all the time to maintain the motivation. In this situation, it is very difficult to maintain praise as genuine and sincere. In contrast, the use of feedback improves performance when it is focused on what needs to be done to improve, and particularly when it gives specific details about *how* to improve.

This suggests that feedback is not the same as formative assessment. Feedback is a necessary first step, but feedback is formative *only if the information fed back to the learner is used by the learner in improving performance*. If the information fed back to the learner is intended to be helpful, but cannot be used by the learner in improving her own performance it is not formative. It is rather like telling an unsuccessful comedian to "be funnier".

As noted above, the quality of feedback is a powerful influence on the way that learners attribute their successes and failures. A series of research studies, carried out by Carol Dweck over twenty years, has shown that different students differ in the whether they regard their success and failures as:

being due to 'internal' factors (such as one's own performance) or 'external' factors (such as getting a lenient or a severe marker);

being due to 'stable' factors (such as one's ability) or 'unstable' factors (such as effort or luck); and

applying globally to everything one undertakes, or related only to the specific activity on which one succeeded or failed.

Table 2 gives some examples of attributions of success and failure.

Attribution	Success	Failure
locus	internal: "I got a good mark because it was a good piece of work"	internal: "I got a low mark because it wasn't a very good piece of work"
	external: "I got a good mark because the teacher likes me"	external: "I got a low mark because the teacher doesn't like me
stability	stable: "I got a good exam-mark because I'm good at that subject"	stable: "I got a bad exam-mark because I'm no good at that subject"
	unstable: "I got a good exam-mark because I was lucky in the questions that came up"	unstable: "I got a bad exam-mark because I hadn't done any revision"
specificity	specific: "I'm good at that but that's the only thing I'm good at"	specific: "I'm no good at that but I'm good at everything else"
	global: "I'm good at that means I'll be good at everything"	global: "I'm useless at everything"

Table 2: dimensions of attributions of success and failure

Dweck and others have found that boys are more likely to attribute their successes to stable causes (such as ability), and their failures to unstable causes (such as lack of effort and bad luck). This would certainly explain the high degree of confidence with which boys approach examinations for which they are completely unprepared. More controversially, the same research suggests that girls attribute their successes to unstable causes (such as effort) and their failures to stable causes (such as lack of ability), leading to what has been termed 'learned helplessness'.

More recent work in this area suggests that what matters more, in terms of motivation, is whether students see ability as fixed or incremental. Students who believe that ability is fixed will see any piece of work that they are given as a chance either to re-affirm their ability, or to be 'shown-up'. If they are confident in their ability to achieve what is asked of them, then they will attempt the task. However, if their confidence in their ability to carry out their task is low, then they will avoid the challenge, and this can be seen in mathematics classrooms up and down the country every day. Taking all things into account, a large number of students decide that they would rather be thought lazy than stupid, and refuse to engage with the task, and this is a direct consequence of the belief that ability is fixed. In contrast, those who see ability as incremental see all challenges as chances to learn—to get cleverer—and therefore in the face of failure will try harder. What is perhaps most important here is that these views of ability are generally not global—the same students often believe that ability in schoolwork is fixed, while at the same time believe that ability in athletics is incremental, in that the more one trains, the more one's ability increases. What we therefore need to do is to ensure that the feedback we give students supports a view of ability as incremental rather than fixed.

Summary

Perhaps surprisingly for educational research, the research on feedback paints a remarkably coherent picture. Feedback to learners should focus on what they need to do to improve, rather than on how well they have done, and should avoid comparison with others. Students who are used to having every piece of work graded will resist this, wanting to know whether a particular piece of work is good or not, and in some cases, depending on the situation, the teacher may need to go along with this. In the long term, however, we should aim to reduce the amount of ego-involving feedback we give to learners (and with new entrants to the school, perhaps not begin the process at all!), and focus on the student's learning needs. Furthermore, feedback should not just tell students to work harder or be 'more systematic'—the feedback should contain a recipe for future action, otherwise it is not formative. Finally, feedback should be designed so as to lead all students to believe that ability—even in mathematics—is incremental. In other words the more we 'train' at mathematics, the clever we get.

Although there is a clear set of priorities for the development of feedback, there is no 'one right way' to do this. The feedback routines in each class will need to be thoroughly integrated into the daily work of the class, and so it will look slightly different in every classroom. This means that no-one can tell teachers how this should be done—it will be a matter for each teacher to work out a way of incorporating some of these ideas into her or his own practice. However, the size of the effects found in the experiments discussed above, and in the other research reviewed by Black and Wiliam suggests that changing the kinds of feedback we use in mathematics classrooms could have more effect than all the government initiatives put together.

References

Black, P. J. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, **5**(1), 7-73.

Brophy, J. (1981) Teacher praise: a functional analysis. *Review of Educational Research* **51** (1) 5-32.

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** 1-14

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.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, **41**(10), 1040-1048.

Good, T. L. and Grouws, D. A. (1975). Process-product relationships in fourth grade mathematics classrooms. Report for National Institute of Education, Columbia, MO: University of Missouri (report no NE-G-00-0-0123).

Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, **119**(2), 254-284.