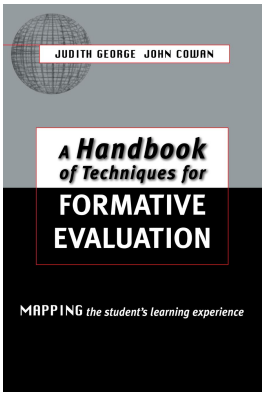
This article was downloaded by: *10.3.98.104*
On: 23 Jun 2021
Access details: *subscription number*
Publisher: *Routledge*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



A Handbook of Techniques for Formative Evaluation Mapping the student's learning experience

Judith George, John Cowan

Choosing a Method of Formative Evaluation – and Using It

Publication details
https://www.routledgehandbooks.com/doi/10.4324/9780203969175.ch0003
Judith George, John Cowan
Published online on: 01 Feb 2016

How to cite :- Judith George, John Cowan. 01 Feb 2016, *Choosing a Method of Formative Evaluation – and Using It from:* A Handbook of Techniques for Formative Evaluation, Mapping the student's learning experience Routledge
Accessed on: 23 Jun 2021
https://www.routledgehandbooks.com/doi/10.4324/9780203969175.ch0003

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

CHAPTER 2

CHOOSING A METHOD OF FORMATIVE EVALUATION – AND USING IT

WHAT DO YOU WANT TO KNOW?

The starting point for your choice is, of course, what you want to evaluate. You may have in mind a very broad area; for example, your students' learning experience in general on your course, or their reactions to your teaching style. On the other hand, your query may be a focused and specific one about the course, the learning environment or your support, the effectiveness of the overheads you used in a particular lecture, the perceived welcome of the introductory session, or the way in which you introduced a new and challenging concept.

A caveat: your choice of a focus could be a two-stage process

On occasion, you may not be able to decide immediately on the exact focus of your enquiry. For example, you might have learnt that the use of CAL drill packages has significantly enhanced learning in a similar course in another institution, and in your own course you judge that there is clear scope for improvement; or you could have seen evidence in examination performance that certain topics are less well grasped than others; or you may have had adverse feedback at a staff–student committee meeting on the tutorial provision. In the first example, you will be well advised to try another approach initially, and then carry out a comparative evaluation – once the new option has been debugged and is reasonably well established. In the last two examples, you will need more information about the status quo before planning changes, and

so will probably want to investigate and iterate towards improvement, by using some method of formative evaluation within the existing provision.

In such situations, it is probably wise to identify and assess any presuppositions which you, or others, are making about the cause of the weakness you wish to minimize or eradicate. Indeed, your first step in any formative evaluation may well be akin to deciding, in the broadest terms, where to aim the telescope of enquiry. If examination performances in your part of a course imply weak learning, should your formative evaluation concentrate on the teaching, the support for student activity, the textbook, the style of examination question, the lack of preparation for examination in that format, or the conflicting demands and attractions of other options in the examination paper? Without further information, you cannot make a defensible choice of focus. A preliminary enquiry will help you to decide where to look more closely, and what data you need to acquire about that aspect of the teaching, learning and assessment. Your formative evaluation will thus be in two stages – first a broad frame enquiry, to determine where to aim a more focused study later; then the focused study.

There is another sound, but different, reason for undertaking an enquiry in two stages. Through no lack of competence or experience, you may need a preliminary run to fine-tune the method of enquiry itself before applying it to the range of subjects or situations from which you wish to obtain data. For example, a recent evaluation was prompted by complaints from students who maintained that the assistants who provided tutorial support did not explain clearly when they were asked for help. An initial set of interviews revealed the interesting outcome that only the poorer students in the class group reported difficulty in comprehending the explanations which they sought and were given; their assertion (understandably) was that the assistants couldn't explain to students what they had to do. Amplification of the answers to follow-up questions asked by the evaluators revealed that poorer students were dissatisfied with the responses from assistants when they (the students) were asking what to do, in the expectation that the assistants would, in effect, do the thinking for the students. In this case the two-stage evaluation process was necessary to permit refinement of the methodology, to concentrate on students' expectations of teaching and their interpretation of course aims.

Be prepared, then, for formative evaluation to take place in two stages, and to be more useful as a result.

THE PURPOSE SHOULD DETERMINE THE METHOD

Many departments and course teams, when asked about evaluation during teaching quality assessment or quality audit, will respond confidently that their

university has a standard questionnaire which is used at the conclusion of every module. The inference is that the use of a questionnaire meets the need for formative evaluation and, indeed, for summative evaluation into the bargain. Yet that is demonstrably unsound reasoning, for a questionnaire, like any other method of formative evaluation, covers only part of the range of matters which merit regular reconsideration.

One of the writers introduced tape–slide tuition to his undergraduate course at a time when that approach was relatively new. His students warmly praised the method, and would have confirmed that verdict if given a questionnaire to evaluate their reactions. They asked for more of this type of teaching – because they had 'learnt so much from it, so easily'. He then asked them to take a simple 90/90 post-test to test their learning anonymously. The result should have been that 90 per cent of the students scored at least 90 per cent on the factual recall test. This was far from the case, to the students' frank surprise. The moral of this little tale is that learning should be evaluated by finding out what students have *actually* learnt, not by what they *think* they have learnt. Questionnaires can surely only elicit factual information of which students have direct knowledge, such as the number of hours of study they put in during a typical week. Otherwise questionnaires obtain opinions. (Questionnaires may even only elicit opinions about the number of hours devoted to study!) Note, however, that the post-test which ascertained learning from the tape–slide sequence did not determine student *reaction* to the innovation, which was also important. The determination of the purpose of enquiry should, then, be matched by the choice of an appropriate method of evaluation; and vice versa.

In broad terms, the purposes that we may have in deciding to undertake some formative evaluation could be categorized under at least four distinct headings. Some methods, such as post-testing or the analysis of students' work, can tell us where *learning* has, and has not, taken place. Others, such as observations, briefed or unbriefed, and the taking of protocols,[1] can tell us about the *learning experience,* rather than the learning outcomes. Yet other methods, such as interpersonal process recall,[2] concentrate on feelings and reactions during learning, that is, the *immediate reactions* to the learning experience. And a final group, as we are choosing to divide them, may inform us of the *values* which students place upon the learning or the learning experience, for whatever reasons – as revealed, for instance, by the writing of letters from students to their successors, in the next academic year.[3]

We, therefore, suggest to you that you should plan evaluation by choosing a method or methods appropriate to what you or your course team want to know rather than, as is often the case, choosing a method for no relevant reason, with the outcomes which follow in consequence, appropriately or otherwise.

THE POTENTIAL TO BRING ABOUT CHANGE IS AN IMPORTANT CONSIDERATION

A detailed study of the effects of modularization on assessment loading for students and staff, and on the assessed coverage of the syllabus, was conducted in a department whose institution was firmly committed to modularization, and unlikely to make any changes in that policy. The analysis of the findings, and the presentation of them, took place in fora where the real agenda was to question the desirability of the *fait accompli* of modularization, despite the futility of that discussion. In a second and somewhat similar department, a comparable evaluation accepted the inevitable and set out to discover the aspects of learning that were suffering, and those that were not the subject of any deleterious effects. The result of that truly formative evaluation was to devise revised teaching and learning strategies and assessment methods to cope to an improved extent with what were perceived and accepted as the *de facto* constraints.

There is no point in engaging in formative evaluation that would focus on the need for changes in the unchangeable. Conversely, if we are in the business of bringing about change, it can be sound strategy, and far from devious, to engage in formative evaluation which should generate findings which will directly *influence* decision-making, as well as *inform* it. We instance an evaluation which confirmed the effect on learning of frequent testing, in a limited allocation of time, thus encouraging shallow rather than deep approaches. That study informed the decision-makers, but did not actually lead to any institutional change. However, when the next evaluation of the same situation followed a different tactic and showed (and publicized) higher retention scores in the modules where testing was less frequent and time pressures less acute, the majority of the departmental staff chose to modify their assessment practice and policy. Choice of focus for evaluation can strengthen its impact or render the effort pointless.

MULTIPLE PERSPECTIVES CAN ENHANCE EVALUATION, AND ITS USEFULNESS

Consider the possibility and potential of a situation in which three approaches to evaluation were adopted – and were chosen to concentrate on rather different aspects of the curriculum. The tutor, for example, arranged for:

1. Dynamic lists of questions,[4] which offered a proven means of judging the planning of class sessions and judging their perceived effectiveness. *(Primary purpose: to determine the students' perception of the effectiveness of these sessions in dealing with their perceived and declared learning needs)*;

2. Blind second marking, by the tutor and others, of the students' submitted work, to a carefully formulated and agreed marking schedule.[5] *(Primary purpose: to confirm the objectivity of marking, and identify any scope for improvement in it)*;
3. A closing activity in class, making it possible for students to formulate advice to the tutor for next year on a 'stop/start/continue' basis – listing what not to do again next year, what to introduce next year, and what to retain for its strength.[6] *(Primary purpose: to identify scope for improvements worthy of consideration, and ongoing strengths that should be retained in the provision)*.

Among the outcomes of the evaluation were the following interrelated findings or suggestions:

- From (1), it was clear that outstanding questions on the dynamic lists were often in the form 'What other points of view should I know about, and consider?'
- From (2), although it was not the immediate focus of enquiry, it emerged that sections of the students' work which often attracted low, and unreliable, marks by assessors (including the main assessor when there was repeat marking) were associated with the students not taking a balanced view, in which optional possibilities and interpretations were properly considered.
- From (3), students were advising the tutor to stop asking them what *they* (the students) thought, and to start telling them more about what authorities in the field had made of such questions (shades of Perry!).[7]

While (1) and (3) had conveyed something of the same message, the findings in (2) pointed to a woolliness in the criteria and marking schedule, and even in the tutor's own thinking about that aspect of the learning. All of this, *taken together,* pointed to the need for the tutor to:

- think through the criteria associated with critical thinking and balanced review;
- communicate these criteria and goals more clearly to students;
- proclaim the criteria explicitly in marking schedules and in formative marking;
- structure part of the tutorials around the development of the required ability, probably with some provision, for example, for reflection on the process.

Notice how the multiple perspectives of a combined trio of approaches led to a deeper and more sensitive identification of, and response to, the need for a particular development in the teaching and assessment combined.

BEWARE YOUR OWN ASSUMPTIONS

- Watch out, for example, for your possible reliance on students' 'letters'[8] or wash-up reports, or other evaluations, which concentrate on aspects of the course which *you* have already made clear to them are important to *you* – such as developing critical thinking, and the accessibility of the tutor. Make sure you also find out about other aspects of the course, such as understanding of key concepts, and the clarity of the pre-course documentation, which you have been less energetic in emphasizing or providing. Adopt the approach that Karl Popper would advocate: concentrate on what may have been neglected, or not considered.
- Watch out for evaluations that concentrate on what you think is worth evaluating (which is not quite the same as the previous point). Make some provision for some other (external) source of focus for evaluations. It can be one of the most important outcomes of questionnaires, for instance, to suggest questions that need to be pursued, and have not so far received attention.
- Watch out for evaluating against criteria which you have not made explicit, and so are not scrutinizing as a matter of course. We recall an instance where the implicit criteria set such a high store on dealing with *dis*advantage (of various forms) that fairness in the treatment offered to the *non*-disadvantaged was losing out. That was not apparent until the criteria were made explicit, and were applied across the board of both tutor and course performance.

In such cases, the best response, once an evaluative weakness is identified, is usually to widen and strengthen your strategy, rather than to discard it.

CONSIDER RESOURCING CAREFULLY – ESPECIALLY IN TERMS OF HUMAN RESOURCE

We make this point through a few disconnected examples:

1. One of us learnt more about his students' learning, and made more changes in his teaching, as a result of taking and analysing recorded protocols[9] than from any other method of formative evaluation. But the price was one that he was not prepared to go on paying. Transcribing protocols is a dreary and difficult business. The subjects do not speak in carefully prepared or considered sentences, they often exclaim or mutter, and the tone and volume change continually. Despite the powerful

usefulness of recorded and transcribed protocols, that writer abandoned them as a feasible option, and sought second-best options with similar but rather less rich results.

2. Questionnaires are useful, but questionnaire surveys often lead to low return rates, because the students are left to return the forms at their leisure, and many do not do so. The result is often a statistically unreliable sample, which comes from a minority with particular reasons for responding – perhaps to log complaints, perhaps to praise. Responding in the students' own time is often accepted for expediency, because response time (in human resource terms) is judged something that cannot be spared in class to issue forms, and have them completed and returned. That is an unfortunate decision, and can have misleading outcomes. We have seen strongly worded evaluations – and consequent decisions – in course records, where the returns came from less than 20 per cent of the class group!
3. Interpersonal process recall,[10] in our experience, is often genuinely (and not evasively) discarded because it is seen as labour-intensive. It calls for perhaps two hours' time from a colleague, and 30 minutes each from two students and the tutor, to lead to outcomes to test at a later date against class opinion – in perhaps 15 minutes of class time. *Yet you would only do all this just once a year, or every other year.* Subjective decision-making notes the durations of effort, but not the frequency, but should balance the weighting in time commitment against the richness of the data acquired by this method. It is important to be sound in the judging of the demands on human resource, in terms of time.

IS THE EVALUATION LIKELY TO CONVINCE THOSE WHO RECEIVE IT?

We begin our answer from our own personal experience. We frequently recall that our early experiences of interpersonal process recall were shattering revelations – of important aspects of our teaching, and of its influence and impact on our learners of which we had hitherto been unaware. Yet it has been our experience, when frankly recounting examples of this to our peers during staff development activities, that they are somewhat dubious about the importance to our students of the findings which we narrate, and that they cannot quite see the need for us to have reacted as we did. We did not, we add, encounter that reaction from the learners who were to profit from the revelations, and our constructive reactions to them. The judgement of relevance, or of the adequacy of rigour in the findings of evaluations, depends on who makes the judgement, and how they judge satisfaction.

Notice, though, that formative evaluation can often be at its most effective when it does no more than suggest neglected aspects of the process which merit remedial or developmental attention. Observations of student reaction and behaviour may reveal lack of understanding, or may note questions from learners who are desperately seeking clarification about the subject, or may record actions at variance with the teacher's declared intentions. They may not rigorously establish *what* is wrong; but they are more than adequate to prompt a review and revision of the teaching and learning situation by showing that all is not well.

Always remember, though, and perhaps be reassured by the view, that there is a great risk in the worship of quantification, and its apparent rigour. Someone (at least one) has maintained sensibly that the things which matter cannot be measured, and the things which can be measured don't really matter.

THE PROCESS

Before formative evaluation begins

Let's just mention an obvious preliminary step. It should almost go without saying that you need to agree with the students a clear definition of the remit for the evaluation, even if and when you yourself evaluate, and a clear understanding of the role of anyone who assists you in that. If you are obtaining, or making public even within the class, information that is normally private or confidential, you must obtain permission to do so.

Analysing the data[11]

The important points to be made under this heading stem from the need for you to do all that should be done to ensure objectivity, and to avoid skipping hurriedly and carelessly through an important stage in the process.

You should separate the assembling of the data from the analysing of it. In the analysis, beware of the temptation to merely summarize and present that summary in a businesslike way. Look for patterns, and for inconsistencies – and point them out once they are found.

When you move on to interpret the data, and decide how to respond, be explicit – to yourself, your students and perhaps a helpful peer – about the values against which your decisions are then made as well as the actions you will take.

Think tactically – in terms of probable acceptability

If you first give thought to the way in which the outcome you hope for is likely to be received, this may suggest ways in which you would use or present your findings. Let's briefly consider a range of possibilities under this heading. You may wish to:

- *Reinforce a need for change*: one of us taught in a department where all the staff were agreed that the teaching of mathematics in schools had deteriorated dreadfully, and that many first year students were incapable of handling basic trigonometry. He eventually reached, and tried to confirm, a counter-hypothesis that it might be the inability of his students to deal with diagrams containing redundant lines, rather than an inability to handle basic trigonometric computations, which required remedial attention. He devised a test to ascertain the grasp of both abilities. He confirmed that the weakness on which remedial tuition was focusing was not present, while the suspected weakness was clearly present. The effort in remedial teaching then changed significantly, as a consequence of that finding – as did its effectiveness.

What was required here from formative evaluation was confirmation or otherwise of a view already formed, but not proven.

- *Amplify a suspicion*: we recall several situations in which we have suspected a need for change, and have wished formative evaluation to confirm this belief, and (in that case) to suggest means of achieving development. For example, observations of student behaviour during certain formal examinations, coupled with analysis of the marks scored for first, second, third and so on questions attempted, led one of us to suspect that students probably did not profit from being offered freedom of choice in the questions which they would answer. Conversely, the conclusion was that it would be worth while to experiment with situations in which students had no freedom of choice in examinations. A tentative investigation on these lines[12] confirmed that offering freedom of choice to these students, in these examinations, was not to their advantage.

This formative evaluation provided confirmation of a suspicion, and led to action. It was undertaken in a situation which called for more strength and authority than would have been the case in response to a mere suspicion in which there was little initial confidence – even on the part of the investigator.

- *Inform review and debate*: a formative evaluation of the interactions and the reactions of students during audio-conference calls opened up the whole possibility that affective, rather than cognitive, outcomes might be the most important for distant student learners in such situations.[13] That, in turn, led to consideration of the implications for telephone-conference tutorial design in terms of changed objectives and methods for the sessions.

Here there was no question of confirming a suspicion, nor was there a conclusion as such; there was merely the production of objective, and relevant, data for consideration, and subsequent action.

- *Discover unperceived needs*: we have already mentioned that the findings from interpersonal process recall generally lead to some shocked surprise on the part of tutors, and even students. This occurs when it emerges that there is a stark mismatch between the students' and tutor's perceptions of their interactions. But, as we hope we have already exemplified, it is usually a shock or surprise received by someone who is open to receive and to consider this type of feedback. In such circumstances, it is important that the reactions drawn from student subjects are honest, factually reported, and above all presented and considered in circumstances which are not threatening or embarrassing for the tutor. For in such cases the issue of the acceptability of the findings is not a real problem, in our experience.

- *Establish an unperceived need*: in this category, we consider problems in formative evaluation involving the work of those teachers who are *not* open to entertain doubts or reservations about their teaching. An evaluation must present such people with incontrovertible findings whose accuracy and relevance are without question. Here the evaluator should be in no doubt about the critical point to be made, while the confident teacher is equally in no doubt that all is well. The difficulty is to find a way of presenting data in order to make the point. For example, we recall the formative evaluation which revealed that, although the terminal examinations contained demanding questions and demanding part-questions, few marks were scored in the face of higher-level cognitive demands – even by the more able students. Marks were accrued by responding to questions and part-questions in which the demand was mundane, and often called for no more than regurgitation. When this was displayed in a form which made the point clearly, the unavoidable conclusion – for all, including the person concerned – was that no students displayed much competence in the objectives to which that teacher had previously claimed high commitment, and – by inference – teaching success.

What was urgently required here was to establish a hitherto unperceived need for improvement.

- *Change attitudes*: in another example of unperceived need, a teacher whose examinations consistently led to the failing of the majority of the students, maintained firmly at exam board meetings that his students were stupid, and that the marks in the other subjects must have been the result of soft marking or simple questions. A careful analysis was made of the performances of students in the examination in question. This showed that the questions appeared, from the student performance in answering them, to be of comparable difficulty, and that, while few students completed all of the stipulated questions, most students scored about 75 per cent of the available marks in what they had managed to attempt. The clear inference was that it was shortage of time, and not of ability, which was the root cause of low marks. The assertion of stupidity fell, although it was then open to the teacher to adopt a new hypothesis, of a slow rate of working! And the redesign of the examinations was established as an urgent need.

The role of this formative evaluation was not only to establish a need that had not been recognized as such, but to do so in a sufficiently powerful way so as to change the attitudes and assumptions underlying previous conclusions.

OUR GROUPING OF METHODS

In the following five chapters, we have collected together methods that have been of use to us and may be of use to you. Our grouping is to some extent arbitrary and personal. We have assembled ideas according to the purpose for which the method is, in our own experience, most useful; and since we recognize that some methods are useful for more than one purpose, we know that this subdivision may create difficulties for some readers, especially those with experience in this field.

Nonetheless, we group together methods that *primarily* obtain information about:

1. the immediate learning experience;
2. students' reactions that occurred during the learning experience;
3. the success of learners in achieving the intended learning outcomes;
4. student reactions *after* the experience.

We have deliberately left until third in our sequence those methods that inform teachers about the success of learners in achieving the intended learning outcomes. Our placing of the immediate learning experience as our first concern itself testifies to the importance that we place, as teachers, on the learning experiences that we create for our students. The next chapter, then, is followed by a chapter describing a linked group of methods, which enable us to obtain information about students' reactions during the learning experience. Only then do we move on to achievement of the learning outcomes, after which we include a section on a somewhat neglected type of data for formative evaluation, which is concerned with the student reactions *after* the experience – and perhaps long after it.

Conclusion

We hope that we have shown that:

- it will usually be best to use several methods of formative evaluation to provide a composite impression of the learning, teaching and assessment;
- methods should be chosen to provide the type of information that the evaluation requires;
- other factors, such as the way the information will be used, and by whom, are also of importance in the choice of method(s);
- it is desirable to give some thought to the values against which judgements will be made from evaluations;
- not all evaluations call for separate activity to generate data; some may use data that are already available, such as portfolios, learning journals and examination scripts.

Notes

1. See Method 3.5, pages 44 and 119.
2. See Method 4.2, pages 60 and 123.
3. See Method 6.4, page 84.
4. See Method 3.2, pages 39 and 115.
5. As discussed in Chapter 8, Question 10.
6. See Method 6.9, page 90.
7. See Perry, W (1970) *Forms of Intellectual and Ethical Development during the College Years: A scheme*, Holt, Rinehart and Winston, New York, who identified nine stages of development; see also Belenky, M F *et al* (1986) *Women's Ways of Knowing: The development of self, voice and mind*, Basic Books, New York.
8. See Method 6.4, page 84.

9. See Method 3.4, page 43.
10. See Method 4.2, pages 60 and 123.
11. For advice on analysis of data in general, see Robson, C (1993) *Real World Research: A resource for social scientists and practitioner-researchers*, pp 303–408, Blackwell, Oxford.
12. Cowan, J (1972) Is freedom of choice in examinations such an advantage?, *Technical Journal*, **10** (1), pp 31–32.
13. See Lee, M (1997) *Telephone tuition project report*, Open University in Scotland, Edinburgh.