

ASSESSMENT – SUMMATIVE AND FORMATIVE – SOME THEORETICAL REFLECTIONS

by MADDALENA TARAS, *University of Sunderland*

ABSTRACT: This paper wishes to clarify the definitions of the central terms relating to assessment. It argues that all assessment begins with summative assessment (which is a judgement) and that formative assessment is in fact summative assessment plus feedback which is used by the learner.

Keywords: assessment, summative, formative, definitions, theory

1. INTRODUCTION

Assessment is of central importance in education, and yet there is a lack of commonality in the definition of the terminology relating to it. Development of both theoretical and practical applications will suffer unless there is coherence and agreement in the definition of the terms. Assessment for learning or formative assessment (FA) is increasingly being emphasised, yet its relationship to summative assessment (SA) has been little explored.

This paper will examine, first, the central terminology and definitions relating to assessment; second, the literature and current theoretical discussions pertaining to summative assessment and formative assessment; and third, implications for the realignment of definitions and practice. It returns to basic principles of SA and FA; that is, to Scriven who first made the distinction.

Since Scriven, there have been developments, both theoretical and practical in the area of assessment. However, the tenets which describe the basis of assessment remain essentially the same.

2. DEFINITIONS

The Terms ‘Evaluation’ or ‘Assessment’

The terms ‘Evaluation’ or ‘Assessment’ are both in common usage in educational circles. In the UK, the common use of the term

‘assessment’ is to refer to judgements of students’ work, and ‘evaluation’ to refer to judgements regarding courses or course delivery, or the process of making of such judgements. This is how these terms will be used in this paper, for the sake of clarity and consistency, other than in quoting writers who have done otherwise. For example, Scriven uses ‘evaluation’, as do most US academics, in the sense of ‘assessment’ as used in this paper.

Definition of Assessment

Like Scriven, I take ‘assessment’ to refer to a judgement which can be justified according to specific weighted set goals, yielding either comparative or numerical ratings. For him it is necessary to justify (a) the data-gathering instruments or criteria, (b) the weightings and (c) the selection of goals (Scriven, 1967, p. 40). I argue that it is necessary to add a further stage – that of justifying the judgement against the stated goals and criteria.

The Assessment Process

The process of assessment is the mechanics or steps required to effectuate a judgement. A judgement cannot be made within a vacuum, therefore points of comparison, i.e. standards and goals, are necessary. The criteria narrow the choices of specific items which are considered important and relevant for any specific judgement within any given context. Therefore, within the process are manifest the parameters within which the judgement is made. In other words, during the process of making a judgement, all these elements are in constant inter-play. All assessments require these parameters and these can either be explicit or implicit (Black and Wiliam, 1998, p. 8; Sadler, 1998, p. 2). With implicit parameters, any second assessor is playing a guessing game as to which points will be salient within an assessor’s mind. Explicit parameters go some way towards creating a shared forum for assessment and therefore allowing transparency of process, although within any given context, meanings can and do vary between individuals (for a discussion of ‘fuzzy criteria’ see Sadler, 1989).

3. SCRIVEN: THE PROCESS OF ASSESSMENT

In his seminal paper, Scriven states that assessment is a single process.

Evaluation is itself a methodological activity which is essentially similar whether we are trying to evaluate coffee machines or

teaching machines, plans for a house or plans for a curriculum. (Scriven, 1967, p. 40)

Therefore, assessment is assessment. Assessment is an integral part of all aspects of daily life. Scriven was working in the context of curriculum evaluation, but the principles he discusses are universally relevant to all assessment.

Summative Assessment

The process of assessment leads to summative assessment, that is, a judgement which encapsulates all the evidence up to a given point. This point is seen as a finality at the point of the judgement. A summative assessment can have various functions which do not impinge on the process.

Formative Assessment

Since the process of assessment is, as Scriven notes, a single process, i.e. making a judgement according to standards, goals and criteria, FA is the same process as SA. In addition, for an assessment to be formative, it requires feedback which indicates the existence of a 'gap' between the actual level of the work being assessed and the required standard. It also requires an indication of how the work can be improved to reach the required standard.

Therefore, both SA and FA are processes. It is possible for assessment to be uniquely summative where the assessment stops at the judgement. However, it is not possible for assessment to be uniquely formative without the summative judgement having preceded it. This SA can be implicit and the formative focus only made explicit, or both SA and FA process can be explicit (Black and Wiliam (1998, p. 8) favour implicit while Sadler (1998, p. 2) requires explicit expression of SA).

The 'Role' or 'Function' of Assessment

'Role' and 'function' are used synonymously in this paper. The functions of assessment refer to its intended use or purpose. The choice of function does not impinge on the actual process of assessment, but it will affect the choices of the parameters of the assessment. Therefore, the goals, standards and criteria will, to some degree, be influenced by the choice of function. Although we would expect the educational functions to predominate, often it is the social needs which come to the fore (Broadfoot, 1996; Filer, 2000).

Functions are many and can be combined into multiple uses (see Black and Wiliam, 1998; Sebatane, 1998 for detailed discussion of this issue). 'Evaluation can and usually should play several roles' (Scriven, 1967, p. 41). The problems then presented by Scriven are very pertinent to the current situation in education. If we wish to prioritise educational goals over social roles of assessment, these truths should be faced and solutions brought into effect.

Failure to make this rather obvious distinction between the roles and goals of evaluation, not necessarily in this terminology, is one of the factors that has led to the dilution of the process of evaluation, to the point where it can no longer serve as a basis for answering the questions which are its goal. This dilution sacrificed goals to roles. (Scriven, 1967, p. 41)

Assessment permeates every aspect of our lives, and is a natural and automatic activity (Rowntree, 1987, p. 4). In the educational context, the terrors evoked by the term 'assessment' have distorted its necessity, centrality and its potentially neutral position. Indeed 'assessment' is considered so negative that the term 'evaluation' was preferred for many years. Currently, FA is the antiseptic version of assessment and SA has come to represent all the negative social aspects.

One of the reasons for the tolerance or indeed encouragement of the confusion between the roles and goals is the well-meaning attempt to allay the anxiety on the part of teachers that the word 'evaluation' precipitates. By stressing the constructive part evaluation may play in nonthreatening activities (roles) we slur over the fact that its goals always include the estimation of merit, worth, value etc., which all too clearly contribute in another role to decisions about promotion and rejection of personnel and courses. (Scriven, 1967, p. 42)

As Broadfoot notes (2000, p. *x*), SA is what is open to abuse. It is this desire to suppress the negative and destructive side effect of assessment which devalues personal worth and future prospects, and which has instigated many educationalists to see SA in a negative light and promote FA (Black, 2003b, c; Black *et al.*, 2004; Sebatane, 1998; Torrance, 1993; Torrance and Pryor, 2001, p. 624; Wiliam, 2000b; Wiliam and Black, 1998). Society at large naturally and rightly, makes judgements; misuse of these judgements does not invalidate or minimise the necessity for the judgements.

One can be against evaluation only if one can show that it is improper to seek an answer to questions about the merit of

educational instruments, which would involve showing that there are *no* legitimate activities (roles) in which these questions can be raised, an extraordinary claim. (Scriven, 1967, p. 41)

It seems that the very fear of the possible social misuse of assessment has distorted our view of it.

Thus there seem to be a number of qualifications that would have to be made before one could accept a statement asserting the greater importance of formative evaluation by comparison with summative. (Scriven, 1967, p. 43)

I part company with Scriven on this point as, like Wiliam and Black (1996) I defend the greater importance of FA over SA in most contexts: firstly, because FA encompasses SA (which I believe should always be explicit); and secondly, FA justifies SA, clarifies how the parameters have been addressed, and what needs to be done.

4. DEFINITION OF FEEDBACK

Before the relationship between SA and FA can be clearly examined and understood, it is first necessary to understand the changes in the theory and definition of feedback.

Ramaprasad

Perhaps the single most important development since Scriven's distinction between SA and FA is Ramaprasad's (1983) definition of formative feedback, which places the theory within the context of complex, multi-criterion judgements.

Feedback is information about the gap between the actual level and the reference level of a system parameter *which is used* to alter the gap in some way. (Ramaprasad, 1983, p. 4, emphasis added)

The definition of feedback adopted in this paper is that proposed by Ramaprasad. It is convenient to call this type of feedback 'formative feedback' (as indeed will transpire in the use of terminology in the literature) since it seems impossible to change the current espoused meaning of 'feedback' from that of 'Knowledge of Results' (see Sadler, 1989).

Although Scriven does not formalise or develop the idea, he precedes Ramaprasad in the belief that formative feedback is integrated into the learning cycle by being used by the learner (see below).

Sadler 1989

Ramaprasad's definition of feedback is closely followed by Sadler's theory of formative assessment in importance. Sadler identifies three conditions for effective feedback, with a key premise for student improvement requiring of students '... the capacity to monitor the quality of their own work during actual production' (Sadler, 1989, p. 119).

[Feedback] requires knowledge of the standard or goal, skills *in making multicriterion comparisons*, and the *development of ways and means for reducing the discrepancy* between what is produced and what is aimed for. (Sadler, 1989, p. 142 emphasis added)

In the above quotation, the first step of the process which makes the qualitative human judgement is SA. This is required to have taken place for there to be feedback; it coincides with Scriven's definition of SA. 'Feedback', for Sadler, is therefore a direct product of SA and is required to be used by the learner; 'formative feedback' is the more usual form of nomenclature to describe this procedure.

The second step is a 'development of ways and means for reducing the discrepancy'. Therefore, formative feedback is targeted feedback to improve learning efficiently and expediently (Sadler, 1989, p. 120).

5. THE RELATIONSHIP BETWEEN FORMATIVE ASSESSMENT AND SUMMATIVE ASSESSMENT

Before examining the relationship between SA and FA and the distinctions made between them in terms of process, it is perhaps necessary to reiterate at this stage that within the educational framework, FA currently refers to two very different contexts. The first is that proposed by Sadler, i.e. FA in complex, multi-criterion contexts (as a learning goal integrated into an achievement goal context). In the second, FA is usually in the context of classroom teaching pedagogy in order to develop learning (as a learning goal). The latter has been the focus of much of Black's work (Black, 2003a, b, c), and that of Wiliam (1994, 2000a, b). I suggest that whether FA is seen as a process assessing a product or a process assessing a process, that it is inherently the same process. FA in all contexts can focus on either or both process and product, thus making this discussion valid in all contexts.

Scriven

Scriven seems to show that the dividing line between SA and FA can be very fine.

Unless entirely ignorant of one's shortcomings as a judge of one's own work, he [sic] is presumably engaged in field-testing the work while it is being developed, and in so doing he gets feedback on the basis of which he again produces revisions; this is of course formative evaluation. (Scriven, 1967, p. 43)

If the field-testing is elaborate, it may amount to summative evaluation of *the early forms* of the new curriculum. (Scriven, 1967, p. 43)

One possible reading of these statements by Scriven would appear to indicate that the difference between SA and FA is a matter of degree of 'elaborateness' (this is how Bloom *et al.* (1971) interpreted the difference and which both Black and Wiliam seem to have followed. Another possible interpretation is that FA focuses on the process of assessing and using feedback, whereas SA tends to focus on the product.

Sadler

Sadler uses Ramaprasad's definition as a basis for his theory of formative assessment. Whether intentionally or not, Sadler's theory conforms to Scriven's principles of assessment. Scriven provides a skeletal framework of assessment, while Sadler's theory of FA develops a complex and coherent rationale. Table 1 provides a brief comparison of how the two processes relate to each other.

I consider it important to add a further stage – that of justifying the judgement in relation to the goals and criteria. This is often necessary in order to explain to learners and other tutors how the assessment has been carried out, therefore creating and building up an informed and shared assessment community.

It is interesting to see how Sadler explains both FA and SA.

Formative assessment is concerned with how judgements about the quality of student responses (performance, pieces, or works) can be used to shape and improve the students' competence by short-circuiting the randomness and inefficiency of trial-and-error learning. (Sadler, 1989, p. 120, emphasis added)

The modal verb 'can' shows that if the judgements are used by the learner there is FA, otherwise, there is just the judgement, which is SA. Therefore, SA, as expected, must come first; it is necessary to assess the quality of the work before feedback can be given for the learner to use. What differentiates SA and FA is that FA is used by the learner. Therefore, according to Sadler, SA does not exclude feedback (or Knowledge of Results).

TABLE 1: Comparative steps for Summative and Formative Assessment

Scriven 1967	Sadler 1989
1. a weighted set of goal scales	1. concept of standards, goals or reference levels
2. gathering and combining of performance data	
3. to yield either comparative or numerical ratings	2. compare actual level with standard
4. in the <i>justification</i> of	
(a) the data-gathering instruments	
(b) the weightings	
(c) the selection of goals	
$1 + 2 + 3 + 4 = SA$	$1 + 2 = SA$
	Feedback possible only after SA
	3. appropriate action to close the gap
	$1 + 2 + 3 = FA$

6. SUMMATIVE ASSESSMENT IS DOWNPLAYED BY SADLER

However, despite the apparent clarity of the relationship between SA, Student Self-Assessment (SSA) and FA in Sadler 1989, it is a clarity which has, to some degree, been imposed by the present author. Sadler mentions SA (and SSA), but does this only in passing. On one level this would seem logical, since his paper presents a theory of FA, but on another, when SA forms the basis of all assessment, it is in fact downplaying the core process.

Sadler minimises links between his theory of FA and the process of SA. In a 20-page document, the following are the only references which link FA with SA. You have to look for them to find them. This may have contributed to the dissociation of FA with SA.

... many principles appropriate to SA are not necessarily transferable to FA; the latter requires a distinctive conceptualisation and technology. (Sadler, 1989, p. 120)

However, from discussions of SA, it is possible to extricate and extrapolate the links between SA and FA.

He notes that the summary grade is likely to detract and draw attention ‘away from fundamental judgements and the criteria for making them’ (p. 121) (which is a very valid comment, see also Taras (2003) and Black and Wiliam (1998)). This, in addition to a further reference, ‘The (macro) process of grading involves the first two ... steps’ (p. 121), gives a clear (and perhaps the only truly explicit

reference to SA) indication that step one and two of his formative theory represent SA.

Further discussions on SA are related to issues of reliability and validity as normally presented in the literature. SA requires reliability (of grades or classification) to take precedence over validity (of assessment). It is beyond the scope of this article to examine the developments in relation to validity and reliability. Therefore, the greatest difference between Scriven and Sadler is that the latter presents SA and FA as having distinctive ideals and processes (although the differences in process are never stated explicitly), while the former accepts different roles, but notes that one inextricably leads to the other.

7. IMPLICATIONS FOR PRACTICE FROM THE REALIGNMENT OF DEFINITIONS AND THE PROCESS OF ASSESSMENT

The clarification of the definitions and processes pertaining to assessment should have considerable impact on practice. To date, FA has been presented as an assessment process as defined in this paper (Black, 2003c, p. 2; Wiliam, 2000b, p. 15), and FA has also been presented as a classroom learning and teaching process (Torrance and Pryor, 2001, p. 624; Wiliam and Black, 1998, p. 8); in both cases, teachers and learners are involved. When FA is considered part of teaching methodology, the teacher is the agent 'doing' something to the learner. In all cases, there is almost total separation of SA and FA.

More recently, Black *et al.* (2004) readjust their ideas on using formative and summative assessment to support each other in the classroom context, essentially because firstly, teachers refuse to separate the two (Black *et al.*, 2004, p. 31) and secondly, because of evidence collected from teachers, who like Taras (2001, p. 610) found the usefulness of using summative assessment for formative purposes (Black *et al.*, 2004, pp. 55–56).

However, even with this development, it means that the current frameworks of assessment require teachers to repeat and duplicate the assessment process if both SA and FA are needed (Black, 2003c, p. 1; Torrance, 1993; Wiliam, 2000b, pp. 1, 15–16; Wiliam and Black, 1996, p. 544). The perceived necessity of duplicating assessment has been prohibitive to the development of FA: teachers, already harassed and overstretched, will not willingly agree to double their workload. Indeed, Black (2003c, p. 1) states that the teachers were persuaded to take on FA as 'extra work'. Furthermore:

Experience in many countries indicates that very few teachers are able or willing to operate parallel assessment systems – one

designed to serve a 'summative' function and one designed to serve a 'formative' function. (Wiliam, 2000b, p. 16)

The truth is happily both more expedient and efficient, and follows the principle of Ockham's razor (Fearn, 2001) which supports the simplest possible solution as being the best.

Therefore, perhaps the most useful consequence of clarifying the relationship between SA and FA, as stated in this paper, is that teachers are no longer required to duplicate an assessment process in order to obtain the information required for both SA and FA. Most SA for formal assessment purposes requires feedback, therefore the only real requirement in order to integrate FA into practice is to engage the learners with using this feedback for learning in future work.

The work of Black and of Wiliam, and that of Torrance and Pryor, and all other researchers and practitioners who promote FA, has been beneficial to teachers because two pedagogic practices conducive to learning have been promoted: firstly, discussing and understanding criteria; and secondly, providing feedback to learners on which they reflect. Also, in practice, some of the problems are attenuated, if not neutralised, by the common sense of teachers when using current processes of FA (Black, 2003c, p. 1; Torrance and Pryor, 2001, p. 622).

[Teacher-Researcher 1] ... said she tried not to make overall judgements, since being in the project, except in combination with more focused critique, but noted that it is impossible to give a critique without at least an implicit judgement. (Torrance and Pryor 2001, p. 625)

This teacher-researcher obviously saw the need of a prior judgement to be made in order to be able to provide feedback.

Another problem exists because the SA process has become implicit and unacknowledged. It is much more difficult to ensure the quality of both process and product for both quality assurance purposes and to ensure shared assessment frameworks if neither are conscious and explicit (Sadler, 1998). An explicit and logical process must surely improve the efficiency of the assessment system, and allow the protagonists to be more conversant and focused in communicating the parameters of assessment.

Perhaps the most detrimental aspect is that FA is seen as a magic formula which is not only separate and distinct from SA, but incompatible with it. This is the perceived 'tension' between SA and FA. Wiliam wishes to save FA from SA:

... we must refuse to accept the incompatibility of SA and FA. Instead we must find ways of mitigating the tension, by whatever means we can. Of course, this is a vast undertaking, and well beyond the scope of this or any other single paper. (Wiliam, 2000b, p. 16)

But this can be done by separating process from function. The complications and tensions which Wiliam evokes would disappear if we follow Scriven. If we accept that in any educational process, assessment requires both SA and FA, then what is done with this information, or what purpose it will serve, is an additional discussion. Therefore, the process is similar for all, the purposes can be multifarious and adapt the information to that (those) purpose(s). The same information, if agreed by the assessors (and those assessed) should be sufficiently clear to be comprehensible even years later (indeed this presumption is inherent in quality control processes). Of course, if different criteria or different standards are used, then the assessment has to commence from the beginning.

Furthermore, by recognising that SA is central and necessary to all assessment, it should stop the demonisation of assessment for validation and certification, and instead see it as a stepping stone to learning. This would be true particularly if FA can be seen as a necessary step which justifies and explains SA.

8. CONCLUSION

To sum up, two issues emanate from Scriven's original distinction between SA and FA. First, Scriven gave us a very powerful tool with the distinction of SA and FA – he permitted us to add the dimension of assessment for learning. He did not wish to create a dichotomy, and clearly indicates that the second dimension of FA can only be in addition to SA for assessment. The power of FA, i.e. assessment for learning, is continuing to be eroded because Scriven's advice has not been heeded and a false separation has been created between SA and FA. The separation has been self-destructive and self-defeating.

Second, Scriven warned against confusing the process of assessment with its socio-educational functions. The focus on its many and myriad functions has compounded the perceived division between SA and FA. SA has been blamed for many problems related to assessment in our education system and as an obstacle to the growth of FA. Until the centrality and indeed neutrality of SA is acknowledged, the real blossoming of FA will not and cannot occur.

9. ACKNOWLEDGEMENTS

I wish to thank John Cowan, my daughter Catherine Angela Jones and Steve Whitley for their support and feedback, and David Hamilton for leading me to Scriven.

10. REFERENCES

- BLACK, P. (2003a) Turning research results into practice: how does the D fit into the R&D? Paper presented at AERA Chicago 21 April. Mathematics Education SIG Session 20.061. <http://www.kcl.ac.uk//depsta/education/hpages/pblackpubs.html>
- BLACK, P. (2003b) (with the King's College London Assessment for Learning Group HARRISON, C., LEE, C., MARSHALL, B. and WILLIAMS, D.) The nature and value of formative assessment for learning. Paper presented at AERA Chicago 22 April, Presidential invited session 34.011. <http://www.kcl.ac.uk//depsta/education/hpages/pblackpubs.html>
- BLACK, P. (2003c) (with the King's College London Assessment for Learning Group HARRISON, C., LEE, C., MARSHALL, B. and WILLIAMS, D.) Formative and summative assessment: can they serve learning together? Paper presented at AERA Chicago 23 April. SIG Classroom Assessment Meeting 52.028. <http://www.kcl.ac.uk//depsta/education/hpages/pblackpubs.html>
- BLACK, P. & WILLIAM, D. (1998) Assessment and classroom learning, *Assessment in Education*, 5 (1), March, 7–74.
- BLACK, P., HARRISON, C., LEE, C., MARSHALL, B. and WILLIAM, D. (2004) *Assessment for Learning: Putting it into Practice* (Maidenhead, Open University Press).
- BLOOM, B.S., HASTINGS, J.T. and MADAUS, G.F. (Eds) (1971) *Handbook on the Formative and Summative Evaluation of Student Learning* (New York, McGraw-Hill).
- BROADFOOT, P.M. (1996) *Education, Assessment, and Society: a Sociological Analysis* (Buckingham, Open University Press).
- BROADFOOT, P. Preface. In A. FILER (Ed.) (2000) *Assessment: Social Practice and Social Product* (London, New York, RoutledgeFalmer).
- FEARN, N. (2001) *Zeno and the Tortoise: how to think like a Philosopher* (London, Atlantic Books).
- FILER, A. (Ed.) (2000) *Assessment: Social Practice and Social Product* (London, New York, RoutledgeFalmer).
- RAMAPRASAD, A. (1983) On the definition of feedback, *Behavioural Science*, 28, 4–13.
- ROWNTREE, D. (1987) *Assessing Students: How Shall We Know Them?* (London, Harper and Row).
- SADLER, D.R. (1989) Formative assessment and the design of instructional systems, *Instructional Science*, 18, 145–165.
- SADLER, D.R. (1998) Formative Assessment: revisiting the territory, *Assessment in Education: Principles, Policy and Practice*, 5 (1), 77–85.
- SCRIVEN, M. (1967) The methodology of evaluation. In R. TYLER, R. GAGNE and M. SCRIVEN (1967) *Perspectives on Curriculum Evaluation* (AERA Monograph Series – Curriculum Evaluation) (Chicago, Rand McNally and Co).
- SEBATANE, E.M. (1998) Assessment and classroom learning: a response to Black and William, *Assessment in Education: Principles, Policy and Practice*, 5 (1), 123–130.
- TARAS, M. (2001) The use of tutor feedback and student self-assessment in summative assessment tasks: towards transparency for students and for tutors, *Assessment and Evaluation in Higher Education*, 26 (6), 605–614.

- TARAS, M. (2003) To feedback or not to feedback in student self-assessment, *Assessment and Evaluation in Higher Education*, 28 (5), 549–565.
- TORRANCE, H. (1993) Formative assessment: some theoretical problems and empirical questions, *Cambridge Journal of Education*, 23, 333–343.
- TORRANCE, H. and PRYOR, J. (2001) Developing formative assessment in the classroom: using action research to explore and modify theory, *British Educational Research Journal*, 27 (5), 615–631.
- WILIAM, D. (1994) Towards a philosophy for educational assessment. Update on a paper given at the British Educational research Association's 20th annual conference, Oxford. <http://www.kcl.ac.uk//depsta/education/hpages/dwliam.html>
- WILIAM, D. (2000a) Integrating formative and summative functions of assessment. Paper presented to Working group 10 of the International Congress on Mathematics Education, Makuhari, Tokyo, August. <http://www.kcl.ac.uk//depsta/education/hpages/dwliam.html>
- WILIAM, D. (2000b) Integrating summative and formative functions of assessment. Keynote address to the European Association for Educational Assessment, Prague: Czech Republic, November. <http://www.kcl.ac.uk//depsta/education/hpages/dwliam.html>
- WILIAM, D. and BLACK, P. (1996) Meanings and consequences: a basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22 (5), 537–548.

Correspondence

Maddalena Taras
 School of Education and Lifelong Learning
 University of Sunderland
 Sunderland SR1 3SD
 E-mail: maddalena.taras@sunderland.ac.uk