

What a Study on the Thoughts and Beliefs in Assessments Reveals About Our Thinking on Summative and Formative Assessment

Maddalena Taras¹, Sana A. M. Almansoori^{2*}, Mark S. Davies³

¹Faculty of Education and Society, University of Sunderland, Sunderland SR1 3SD, UK

²College of Engineering, University of Bahrain, Isa Town Campus, Kingdom of Bahrain

³University of Sunderland, Sunderland, SR1 3SD, UK

DOI: [10.36348/jaep.2024.v08i03.006](https://doi.org/10.36348/jaep.2024.v08i03.006)

| Received: 01.02.2024 | Accepted: 09.03.2024 | Published: 27.03.2024

*Corresponding author: Sana A. M. Almansoori

College of Engineering, University of Bahrain, Isa Town Campus, Kingdom of Bahrain

Abstract

Assessment's increasing prominence in educational discourses and practices, especially in the past 20 years, has not always resulted in increasing clarity as to what the new terms in vogue actually mean. Questions include: does carrying out formative assessment really mean that we have to duplicate assessment; is summative assessment inevitably linked to problems and negative perceptions, as the UK's National Student Survey results seem to indicate; are summative and formative assessment different beasts or are they two sides of the same coin as Scriven (1967) noted when he made the original distinction; is assessment really our weakest link? By asking questions to ascertain tutor beliefs and understandings, we clarify the areas where these misunderstandings arise in a Bahraini context. This work builds on research in different contexts which similarly found that the clarity of our assessment understandings is tainted by contradictions in the literature. The questionnaire used here asked neutral questions on assessment and collected both quantitative and qualitative data. Several responses were consistent, all: used formative assessment, associated it with feedback, which appears essential in their perception, and almost all valued theory. However, inconsistencies were in: definitions of formative and summative assessment, their functions and uses, and the relationship between them. How we interpret the results, to some degree, reflects our epistemological and theoretical positions; even though most of the data were unambiguous. By questioning our own beliefs, we found that we are not immune to our emotions and hopes influencing our interpretations.

Keywords: Assessment Literacy, Formative, Staff Development, Summative, Theory.

Copyright © 2024 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY-NC 4.0) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

INTRODUCTION

Assessment is a particularly difficult research topic because of the specific and idiosyncratic nature of assessment protocols and often, terminology. One example is 'assessment literacies', which are themselves idiosyncratic and disparate, though this is only recently coming to be recognized (Price *et al.*, 2012; Willis *et al.*, 2013; Davies and Taras, 2018). Our work has significance because it explores assessment literacies in higher education using a questionnaire about assessment terminologies and their relationships, in an attempt to reveal the underlying theoretical premise and epistemologies across different staff groups (e.g. Taras and Davies, 2013, 2014; Davies and Taras, 2018).

Here we explore the understandings of assessment among academic staff involved in the delivery of development programmes for tutors at a

university in Bahrain. Although this work extends our earlier work in the UK that used the same tool to explore assessment literacies in both tutors (Taras, 2008; Taras and Davies, 2013, 2014) and staff developers (Taras and Davies, 2017; Davies and Taras, 2018) we are not explicitly exploring the Bahraini context, rather our aim is to test understandings in a model group of tutors accessible to us, and to use our own analysis of the data to expose our own understandings, consensual or otherwise. Further we recognize and comment on the value of working cross-culturally that enabled us to question our own and each other's opinions and beliefs. Clarifying this assessment 'baggage' and coordinating our responses to the data has been both arduous and enlightening.

Background

As far back as 1996, Biggs distilled the concept of constructive alignment for course design and

pedagogy, which essentially emphasized the need for coherence between all aspects of theory and practice, supported by empirical research, within assessment, learning and teaching (ALT). Post-millennium, much of this has been related to integrating learning and teaching to the new discourses of learner and learning-centredness.

For assessment, this meant primarily ensuring that student work would follow logically from teaching and learning and not be added to the end of a course. Using formative assessment (FA) during the course to support learning and outcomes was part of aligning ALT. The biggest issue within this happy family was summative assessment (SA), which was ostracised and excluded from these positive discourses (Lau, 2016; Taras, 2012).

The impact of Black and Wiliam (1998), across sectors and educational ages, exacerbated the dichotomy between SA and FA because they marginalized SA from their work; Biggs (1998) warned against this separation. Assessment for Learning (AfL) developed by the Assessment Reform Group and using the work of Black and Wiliam has ironically three out of the four areas of practice (“questioning, feedback through marking, peer and self-assessment, and the formative use of summative tests”, Black *et al.*, 2003 p5) representing AfL, are unambiguously SA, and it could be argued that questioning can also fit into this category. This fact, that AfL is made up of essentially summative assessments of process and product, seems to be lost on many subsequent researchers with the notable exception of Taras (2007, 2009, 2012). Black and Wiliam’s article has since come under closer scrutiny as concerns their arguments and the research selected: Dunn and Mulvenon (2009) argue cogently that Black and Wiliam’s claims for generalisability of the four articles which they use as a basis for their AfL practices are misplaced and unfounded. This means that the four practices are misrepresented both in their description and also in their generalisability.

It can be argued that the original AfL literature has demonstrated much goodwill but less practical benefits for tutors and students. Much of the AfL literature has been based on discourse and rhetoric and Taras has demonstrated that clever use of metaphor and discourses have won over logic and clarity of processes in important aspects (Taras 2007, 2009, 2012).

With Assessment for Learning came its counterpart Assessment of Learning (AoL). This dichotomy was identified and made synonymous with that of SA and FA (Lau, 2016; Taras, 2016). Thinking along either of these dichotomies is both unproductive and unprofitable in terms of time, energy and learning focus for both students and tutors because it requires unnecessary duplication of processes (Taras, 2009, 2012). A pertinent question might be: if this is what we

believe, why in this article do we use the classifications of assessment ‘for’ and ‘of’ learning? The reply is primarily because much thinking in assessment has been directed along these lines and thus it is more expedient on the one hand, and a truer reflection of opinions and debate on the other (however misguided).

Black (2015) mitigates over two decades of negative discourses linked to SA and the irreconcilable differences between SA and FA in the work of Black and Wiliam. Therefore, the clear dichotomy seems on the surface to no longer be the main position. Developing coherent and cogent assessment literacies against this background is difficult for all educators, in whatever sector.

This paper explores the thinking and beliefs of staff developers and lecturers, because although we do not all need to understand things in exactly the same way, there does need to be a shared and sharable understanding of the basic principles and theoretical framework. Otherwise, we will be likely to work at cross purposes and this poses particular difficulties for students who often complain that different lecturers assess and mark work differently.

This paper supports the theoretical framework presented by Taras (2015, 2016), where essentially all assessment is a judgment made against implicit or explicit criteria and standards. All the results of this assessment, a summation at that moment in time, be they letter or number grade, mark or comment; these may be used by learners to improve their work and it is this use that makes the information become formative. Ethical considerations would require explicit sharing and justification of all assessment parameters and results with all incumbents and stakeholders, such as second markers, moderators, parents, etc. This explains the theory and practice behind the process of assessment regardless of context or medium, whether within or outside of education. Within this framework there is no dichotomy between SA and FA, rather, all assessment follows the same procedure and process, and ultimately any information through grades or discourse presents an option for learners to improve their work.

We focus on process of assessment, that is, what staff are doing and why, rather than on functions of assessment, that is, how the assessment is used. This is because functions, which have dominated AfL research and discourses, can neither be controlled nor limited either prior to, during, or post assessment. Process allows flexibility in nuancing the interpretation of criteria and standards, which can be clarified within collaborative and shared forums, and ultimately is what both staff and students need to know. Basing our assessments of students on functions, which may sometimes have its uses, will nevertheless likely cloud the issue of process. Functions, then, are best left out of assessment discussions (Taras, 2012). Or are they? We shall see in

our discussion of the results that functions refuse to go away because they are linked to our hopes and beliefs in an ethical and transparent assessment system.

Assessment Literacy

Assessment literacy as the words indicate, concerns being informed and knowledgeable about assessment. Willis *et al.*, 2013 define it as:

'dynamic social practices which are context dependent and which involve teachers in articulating and negotiating classroom and cultural knowledges with one another and with learners' (Willis *et al.*, 2013 p241).

This is an interesting definition because it encompasses different practices and contexts and importantly requires *'teachers in articulating and negotiating classroom and cultural knowledges with one another and with learners'*. However, this is not the end of the story. Outside of the specific teacher assessment literacies context dealt within this definition, where practice is the focus, there also needs to be a theoretical basis. Theory is important because it provides the rational and coherent relationship between the elements. Theories are rare because few provide coherent, rational relationships with a generic framework: they are necessary for a common epistemological premise to support varied forms of practices.

Xu and Brown (2016) carried out a scoping review of assessment literacy studies by linking *'two fields of research: educational assessment and teacher education'*. First, why are there two fields? Surely educational assessment must be an integral and important part of teacher education, especially since this is such an onerous and central part of the work of teachers. Second, their review (both the trawl of old research and their 'new' 'conceptualisation' of assessment literacy) excludes explicit discussion, explanation and justification of theory, which would provide a cogent definition that can be used in all contexts. Their initial definition is from Stiggins (1991): *'Assessment literacy (AL), traditionally defined as a basic understanding of educational assessment and related skills to apply such knowledge to various measures of student achievement (Stiggins 1991)'* (Xu and Brown 2016 p149)

The problem with this definition is that *'a basic understanding'* and applying *'such knowledge'* are neither related nor similar skills. Also, does *'a basic understanding'* apply to theory, empirical research, or practice? This definition encompasses so many disparate elements that it becomes meaningless. Furthermore, the definition from Stiggins seems to change: *'To help students attain higher levels of academic achievement, teachers need to develop appropriate types and levels of AL'* (Stiggins 1995).

It is no longer about *'a basic understanding'* and how to apply it, but there are also *'appropriate types*

and levels'. What does this mean? The first definition is to *'measure student achievement'* and the second is to *'help students attain higher levels of academic achievement'*, which again are two very different things. The heart-warming conclusion is that whatever you want to do with assessment, *'assessment illiteracy abounds'* (Stiggins 2010 p233).

At the end of the 'reconceptualisation' by Xu and Brown (2016) we are left with a 'to-do' list for teachers but no working or workable definitions of what assessment literacy might be. Understanding our assessment literacies, like everything else, is assessing (i.e. SA) so that further reflection and action (like FA) will improve our understanding. Our working definition in this paper is:

'Assessment literacy (in any given context) may be defined as an understanding of the issues, general and specific criteria, and standards which may enable an individual to communicate efficiently with individuals in a similar context and also to negotiate meaning (coherently) from an informed position, on assessments of processes or products made within that context.' (Taras in progress, in Davies and Taras 2018).

From this, assessment literacy in an educational context requires two aspects, first, expertise in the subject or context being assessed, and second, expertise in assessment parameters and their relationality. The latter is the definition and theory of assessment which is being adopted in the context. In education in general and higher education in particular, theory may not be high on the agenda.

Tight (2004) found that research in education in HE is generally practice-based and lacking in theory: *'...those based in education or higher education departments ...are usually working in what might be called ...fields of study and practice rather than disciplines. ...there are strong pressures on those working in educational departments to focus their research energies on identifying what works best in the classroom or lecture theatre ... the demand for evidence-based practice gives relatively little priority to theory.'* (Tight 2004 p406).

Tight (2004) classifies and categorises 406 articles on HE in 17 HE specialist journals to begin to quantify the theoretical perspectives used. He notes 8 key themes/issues (p397), 8 key methods/methodologies (p397) and 7 levels of analysis (p397). In order to ascertain an initial measure of the level of theoretical explicitness he uses a 3 point-scale (p400): 1. Explicit, 2. Some evidence, 3. No evidence. The results showed 104 (25.6%) used theory explicitly, 66 (16.3%) used some theory and 236 (58.1%) of the articles were a-theoretical. He concludes that in the wider research community this might be seen as a problem:

'I do think, however, that there is a need for more theoretical engagement so that the field (or community

of practice) can develop further, and gain more credibility and respect.' (Tight 2004 p409).

Our work here and in previous work (Taras 2008; Taras and Davies 2013, 2014, 2017; Davies and Taras 2016, 2018) is unique because it explores the theoretical bases and premises of assessment literacies and thus understandings. Assessment, which is a universal and ubiquitous process, would also be expected to have a universal and ubiquitous theory. We can no longer rely on the automaticity of the past where assessment was not seen as an issue. Now it is an issue because so many terms are used that are not shared, whether between tutors and tutors or between tutors and students. We have common words but we do not have a common language with which to use those words.

MATERIALS AND METHODS

University Context

Development initiatives for academic staff in learning and teaching coalesced when the University launched a Post Graduate Certificate in teaching and learning programme in 2006 in collaboration with a UK university. The goal from the outset was to prepare new academic staff for their role in HE and build the capacity of staff at the University to run the programme in-house. The joint programme ran from 2006–2012, producing approximately 118 graduates, who also received professional recognition as Fellows of the UK's Higher Education Academy (HEA). For the capacity building aspect, the involvement of interested graduates of the programme in its delivery was increased gradually over the 6 years. This began with mentoring that evolved into contributions to the preparation and delivery of workshops. In 2013 University staff were sufficiently mature to develop the University's own version of the programme, which was accredited later that year by the HEA and is still in operation.

Participants

Participants, though drawn from across a single University's academic profile, comprised a team assigned to a unit dedicated for the development of Teaching and learning in HE. They were all involved in the development of academic staff in the University by delivering bespoke workshops and by participating in the delivery of programmes in learning and teaching that are accredited by the HEA and lead to its Fellowship awards. Their roles in the delivery of the programmes differed slightly in that some are programme coordinators, some lectures, some mentors, and some a combination. The unit team (the participants) consists of 17 staff, all recognized by the HEA at either Fellowship or Senior Fellowship category. They all joined the unit on a voluntary basis, choosing to take on the development of their colleagues in the University in addition to their usual teaching, research and administrative responsibilities in their respective departments. Given the context, the staff developers will provide the generic

principles and possible processes and the participants will translate these into their own subject contexts.

Questionnaire

All data are from a single, evaluated, questionnaire of 44 questions (Appendix 1), issued to participants over a two-month period. Evaluation and development of the questionnaire is fully discussed in Taras (2008). An electronic version was prepared on Google Forms, the URL link sent to each member of the team with simple instructions on answering the questions. Participants were asked to answer the questions sequentially and not to go back to previously answered questions, and to answer the questionnaire without discussing it with colleagues. Participants took between 15 and 20 minutes to complete the questionnaire (from the time the link was sent until a response was submitted on Google Forms).

Analysis of Responses to Questions Asking for Definitions of FA and SA

The three authors initially decided to carry out the data analysis of the questionnaires over a social media discussion forum so as to present a common, unified voice in our paper. However, we soon realised that our differences seemed irreconcilable at that point. We then made individual interpretations focusing solely on the four possible interpretations of responses to Q1 and Q43, which we tentatively classified as 'of' learning, 'for' learning, 'both' and 'neither'. The results of the discussion in deciding whether rough definitions of SA and FA represent the notions 'of' or 'for' learning exemplified the difficulties we had reaching an agreement. Our initial individual classifications of the rough definitions of SA and FA were discussed as we met again in virtual plenary, where our individual differences in interpretation were exposed. While we were able to agree on most classifications, some remained refractory, and we focused individually on these.

RESULTS

All percentages are reported on the basis of 17 respondents, i.e. all respondents answered all questions. All respondents teach both their colleagues and University students. We did not ask them to separate out their responses in relation to whose learning they were supporting, since we wanted to gauge their general understandings. Thus 'student' as used below and in the questionnaire may refer to any persons taught by the participants, both University students and teaching staff.

Quantitative Responses

All respondents use FA both with their students and in class while 71% of respondents use FA in homework tasks. 71% separate FA and SA and 35% conflate them (Table 1). Two of the 17 respondents answered "Yes" to both separating (Q8) and conflating (Q9) SA and FA tasks while one answered "No" to both questions, which seems to be contradictory. Respondent

8 provides an example of conflating SA and FA in response to (Q10) of “... *let the students correct their midterm exam mistakes in class*” and “... *discuss together their misconceptions and how to avoid them in the future*”.

For questions related to sharing information about FA with students (Table 2), 88% of respondents inform their students that the assessment is formative. Of these 15, 12 respondents (71%) also explain how it is formative to their students. 76% of respondents mark formative assessment work but 47% grade it. One respondent graded the FA without marking and six respondents marked the FA without grading. 82% of respondents think that FA and SA are related (Table 2).

For questions related to the use of self-assessment (Table 3), 82% of respondents ask their students to self-assess. 76% present it as FA while 29% present it as SA. 18% present it as both FA and SA. One respondent answered “No” to Q17, indicating that his/her students do not carry out self-assessment but still answered “Yes” to Q18 (that it was a formative exercise).

94% of respondents believe that theory is important for teachers. There is an agreement on the use of SA for grading with all the respondents using SA for end-of-course grading and 88% making use of SA for mid-course grading (Table 4). However, fewer (47%) use FA for mid-course grading and even fewer (41%) use FA for end-of-course grading. 18% indicated that they use both FA and SA for mid-course and end-of-course grading (Table 5). 47% do not use FA for grading at all and 12% use FA for mid-course grading only. All respondents reported that SA assessed product while 29% also indicated SA assessed process (Table 5). Conversely, 18% reported that FA assessed product and almost all (94%) indicated that FA assessed process. 65% thought that SA assessed only product and FA assessed only process. One of the respondents reported that SA and FA each assess both product and process.

82% of respondents believe that SA is for validation and 88% believe that FA is for learning (Table 6). Fewer (53%) reported that SA is for learning and that FA is for validation (65%). 29% indicated that SA was only for validation and FA was only for learning while 12% thought the opposite. One respondent noted that SA was neither for validation nor learning but that FA was for both. 35% indicated that both SA and FA are used for validation and for learning (Table 7). All respondents reported that FA provided useful feedback while 47% thought that SA provided this (Table 7). 76% thought that SA and FA are different processes and 24% thought they were similar. Almost all (94%) were sure of how FA and SA are related (Table 8).

There seems an almost universal agreement that students both understand and focus on SA (Table 9).

Further, 59% thought that students understood FA and even fewer (24%) thought that students focused on FA. 18% thought that students understood and focused on both FA and SA. One respondent thought students did not understand either but still focused on SA.

Qualitative Responses

Six (35%) respondents reported that they conflated FA and SA tasks (Q9). All examples given (Q10) could be classified as both ‘for’ and ‘of’ learning, except for one where what had been formative work was used summatively to compensate for an otherwise poor summative mark. In this case clearly the tutor had been keeping an ‘informal’ record of students’ performance in the work that was presented as formative. Interestingly, one respondent understood conflation as giving feedback on draft versions of an ultimately summative report. The remaining four all issued feedback to students from summative graded work.

In response to question 16, most (71%) indicated that formative work, in particular feedback, is seen as preparatory to summative work. One specifically noted that feedback leads to success in summative assessments, and one that FA could lead to SA.

Responses to Questions Asking for Definitions of FA and SA

Our initial individual classifications of the rough definitions of SA and FA are given in Table 10 and 11, respectively. When we analysed our responses together, we were eventually able to identify that our differences were based on interference from our including and superimposing functions and/or purposes into our thinking, and making assumptions about what, for example, ‘grading’, ‘marking’ or ‘final’ meant both theoretically and in practice. We came to the conclusion that since all assessments involve a judgement of what learning has, or has not taken place, they can all be classified as ‘of’ learning. The alternative, that ‘of’ learning is associated with grading or validation or occurs at a particular time during or after a course of study, or is undertaken by a particular person, is fraught with difficulties of definition, shaped by individual notions relating to purpose and functions of assessment. Arriving at such a definition of ‘of’ learning that the three of us could agree on was difficult and we contend that agreement across a wider group, such as members of a subject discipline would be equally difficult. Hence, we chose to use the simpler definition. To clarify, ‘of’ can be agreed as providing an assessment or judgement; grade, validation, feedback or any communication of these may be considered an optional extra. For example, grading or producing a grade, whether to validate a piece of work or not is a declaration of the worth of that work and this should be acknowledged. Based on the above

definition, we were able to agree on the classifications as shown in Table 12, though we return to our troubles in the discussion.

Table 13 shows the consistency of classified responses to the repeated questions asking for definitions of SA (Q1 and Q43) and FA (Q3 and Q44). Numbers in columns 3 and 4 refer to frequency of responses that were classified according to the scheme in columns 1 and 2 ('for' learning (for), 'of' learning (of), and neither). For Questions 1 and 43, which asked for a definition of SA, 47% of definitions were consistently 'of' learning to both questions and 24% were initially 'of' learning but the final response was neither. 29% were initially neither but in response to Q43, out of these 5, 3 definitions were 'of' learning, one was 'for' learning and one response remained neither. For Questions 2 and 44 which asked for a definition for FA, 47% of definitions were

consistently 'for' learning to both questions and one response was consistently 'of' learning. 24% were initially neither, of which only one final response became 'for' learning. 12% were 'of' learning for Q2 and 'for' learning for Q44 and 12% were initially 'for' learning but were neither in the final definition.

We examined the responses to the questions asking for definitions of SA and FA semantically (Table 15) and found that over half the respondents associated notions of 'final' or 'end' with SA, and 'on-going' showed some association with FA. The inclusion of a 'grade' or 'mark' was associated, though not strongly, with SA, but was not associated with FA. 'Measure learning' was slightly more associated with FA than SA, and 'feedback' and 'improve learning' were more prevalent in definitions of FA, but were by no means a universal feature. Some respondents defined FA, but not SA, in terms of 'adapting teaching'.

Table 1: Formative assessment tasks used with students

Responses	5. Do you use formative assessment tasks with your students?	6. Do you use formative assessment tasks in class?	7. Do you use formative assessment tasks for homework?	8. Do you keep summative and formative tasks separate?	9. Do you conflate summative and formative tasks? ('conflate' means 'place or use together')
Number of "Yes"	17	17	12	12	6
Number of "No"	0	0	5	5	11
% "Yes"	100	100	71	71	35
% "No"	0	0	29	29	65

Table 2: Information shared with students about FA

Responses	11. Do you tell them it will be a formative assessment?	12. Do you explain how it will be a formative assessment?	13. Is formative work marked?	14. Is formative work graded?	15. Is formative work related to summative work?
Number of "Yes"	15	12	13	8	14
Number of "No"	2	5	4	9	3
% "Yes"	88	71	76	47	82
% "No"	12	29	24	53	18

Table 3: Use of Self-Assessment

Responses	17. Do your students carry out self-assessment?	18. Do you present self-assessment as a formative exercise?	19. Do you present self-assessment as a summative exercise?	20. Does self-assessment use both summative and formative assessment?
Number of "Yes"	14	13	5	9
Number of "No"	3	4	12	8
% "Yes"	82	76	29	53
% "No"	18	24	71	47

Table 4: FA and SA use for grading

Responses	22. Summative assessment can be used for end of course grades.	23. Formative assessment can be used for end of course grades	24. Summative assessment can be used for mid course grades.	25. Formative assessment can be used for mid course grades.
Number of "Yes"	17	7	15	8
Number of "No"	0	10	2	9
% "Yes"	100	41	88	47
% "No"	0	59	12	53

Table 5: FA and SA assess product or process

Responses	26. Summative - assesses product.	27. Summative - assesses process.	28. Formative - assesses product.	29. Formative - assesses process.
Number of "Yes"	17	5	3	16
Number of "No"	0	12	14	1
% "Yes"	100	29	18	94
% "No"	0	71	82	6

Table 6: FA and SA assess for validation or learning

Responses	30. Summative - assesses for validation	31 Summative - assesses for learning.	32. Formative - assesses for validation.	33. Formative - assesses for learning.
Number of "Yes"	14	9	11	15
Number of "No"	3	8	6	2
% "Yes"	82	53	65	88
% "No"	18	47	35	12

Table 7: FA and SA use for feedback

Responses	34. Summative provides useful feedback.	35. Formative provides useful feedback.
Number of "Yes"	8	17
Number of "No"	9	0
% "Yes"	47	100
% "No"	53	0

Table 8: FA and SA process

Responses	36. Summative and formative are different processes.	37. Summative and formative are similar processes.	38. I am sure/not sure how summative and formative relate to each other.
Number of "Yes"	13	4	16 (Sure)
Number of "No"	4	13	1 (Not sure)
% "Yes"	76	24	94
% "No"	24	76	6

Table 9: Student understanding and focus

Responses	39. Students understand summative assessment	40. Students understand formative assessment.	41. Students focus on summative assessment.	42. Students focus on formative assessment.
Number of "Yes"	16	10	17	4
Number of "No"	1	7	0	13
% "Yes"	94	59	100	24
% "No"	6	41	0	76

Table 10: The initial individual responses to the definitions of SA (authors identified by initials). Disagreements shown in bold

Respondent	Q1. Give a rough definition of summative assessment			Q43. Without looking back, give a definition of summative assessment		
	MT	MD	SAM	MT	MD	SAM
1	of	of	of	of	of	of
2	of	of	of?	of	both	neither
3	of	of	of	of	of	of?
4	of	of	of	of	of	of
5	of	of	of	of	of	of
6	of	both	neither	both	for	both
7	of	of	of	of	of	of
8	of	of	neither	of	of	neither
9	of	of	of?	of	of	of?
10	of	of	of	of	of	of?
11	of	of	of	both	of	both
12	of	of	neither	of	of	of
13	of	of	of	of	of	of?
14	of	of	of	of	of	of
15	of	of	of	of	of	of
16	of	of	neither	of	of	of?
17	of	of	of	of	of	of

Table 11: The initial individual responses to the definitions of FA Disagreements shown in bold

Respondent	Q3. Give a rough definition of formative assessment			Q44. Without looking back, give a definition of formative assessment		
	MT	MD	SAM	MT	MD	SAM
1	of	of	of	for	of	for
2	neither	both	neither	neither	both	neither
3	for	for	for	neither	unscorable	for?
4	for	for	for	for	for	for
5	for	for	for	for	for	for
6	neither	both	neither	for	for	for
7	for	for	for	for	for	for
8	for	for	for	for	for	for?
9	for	for	both	for	for	for?
10	neither	unscorable	both	neither	for	for?
11	for	both	both	of	of	for
12	for	for	both?	for	for	for
13	for	for	both	neither	unscorable	for
14	of	both	of	of	both	neither
15	of	of	both	for	for	for
16	for	for	both?	for	for	for
17	neither	unscorable	for	neither	unscorable	for

Table 12: Agreed Classification of the definitions of SA and FA

Respondent	Q1. Give a rough definition of summative assessment	Q43. Without looking back, give a definition of summative assessment	Q3. Give a rough definition of formative assessment	Q44. Without looking back, give a definition of formative assessment
1	of	of	of	for
2	of	neither	neither	neither
3	neither	of	for	neither
4	of	of	for	for
5	of	of	for	for
6	neither	for	neither	for
7	of	of	for	for
8	neither	neither	for	for
9	of	neither	for	for
10	of	neither	neither	neither
11	of	of	for	of
12	neither	of	for	for
13	of	neither	for	neither
14	of	of	of	of
15	of	of	of	for
16	neither	of	for	for
17	of	of	neither	neither

Table 13: Frequency of various permutations in the definitions of SA and FA in the repeated questions

Classified response to question 1 or 3	Classified response to question 43 or 44	Frequency for definitions of SA		Frequency for definitions of FA	
of	of	8	47%	1	6%
for	for	0	0%	8	47%
of	for	0	0%	2	12%
of	neither	4	24%	0	0%
for	neither	0	0%	2	12%
neither	for	1	6%	1	6%
neither	of	3	18%	0	0%
neither	neither	1	6%	3	18%

Table 14: Semantic analysis of questions asking for definitions

words and phrases used	Q1. Give a rough definition of summative assessment	Q43. Without looking back, give a definition of summative assessment	Q3. Give a rough definition of formative assessment	Q44. Without looking back, give a definition of formative assessment
final / end	11	9		
on-going			4	4
grade / mark	11	6		
grading not included			2	3
no grade / mark			3	1
validation		3		
measure learning	3	5	6	8
feedback		2	6	6
no feedback	1	1		
improve learning			5	8
measure teaching		1		1
adapt teaching			4	2

DISCUSSION

It is certainly desirable, if not imperative, that staff involved in faculty development programmes show a consistent attitude towards and shared understanding of “assessment literacy”, however defined; and even defined locally would be preferable over no definition as at least a starting point for debate. Several participant responses were encouragingly consistent: all employed FA, a relatively new notion at this University introduced as compulsory for all coursework in 2012; all engaged with FA in class, where they are present, but less so in homework tasks, where students work alone; all associated FA with the notion of feedback, so feedback appears an essential component of FA in their perception; almost all valued theory.

In relation to students’ understanding of assessment, there is agreement that SA is better understood and is the main focus of students. For FA, most use the terminology ‘formative assessment’ with their students and explain what it is, which indicates an understanding that students need to be involved and aware of the process for FA to actually be formative and lead to learning. However, the endeavours of staff are somewhat in vain since fewer staff believe that students understand FA and even fewer believe that students focus on FA. One view is that the practice of formally using FA, integrated into the delivery is relatively new at the University and takes time to embed. An alternative is that the practice has been formalized for four years at time of data collection and the participants ought to be championing FA among their fellow tutors.

A comparison of the definitions given by respondents at the start and end of the questionnaire shows that 47% defined SA and FA in both definitions as ‘of’ and ‘for’, respectively, which is consistent with the common discourse in assessment. However, 48% changed their definitions of SA, at least by our classification, and 30% changed their definitions of FA. We assume changes are a result of the process of

completing the questionnaire. The questionnaire had equal representation of questions on FA and SA and so we conclude that changes reflect the thinking processes of participants and clearly SA is more mutable than FA. For SA the issue was confusion around whether it was ‘of’ learning or something more nebulous (‘neither’), with almost equal numbers moving in either direction. For FA there was no clear pattern. Also, the questionnaire promoted a re-analysis of both definitions. The semantic analysis revealed that the questionnaire promoted a significant reduction in SA definitions that included ‘mark’ or ‘grade’ and an increase in those that included ‘validation’ and ‘feedback’. More respondents regarded FA as a tool to improve learning in the second definition than the first definition, where there was an almost equal agreement of FA as a tool for adapting teaching and improving learning. Thus, questionnaires such as ours may be tools in their own right to engage tutors in questions of assessment literacy.

The discussion becomes more problematic when examining responses of assessment product and process. There is a clear divide that SA is assessment of product and FA of process, though a minority regard both FA and SA as assessing both. We encountered great difficulty in coming to a consensus on whether definitions supplied by participants were ‘for’ learning, ‘of’ learning, ‘both’ or ‘neither’, and yet we have considerable experience in not only teaching and learning in HE but in its development, management and research. Under these circumstances how can we expect those with less experience to find consensus, unless it is through a simplification of arguments or through simply regurgitating what they themselves have been taught? Further our experiences suggest that the terms ‘for’ and ‘of’ learning may be artificial and unhelpful. The notions of function and/or purpose of assessment are linked to culture, and socio-political stance, experience and pressures. These ideological beliefs are so powerful to us as individuals that they interfere with our ability to arrive at less subjective definitions. These were the pressures

which made it difficult for us to decide what we mean by SA and FA. We were wary of being influenced by the terms ‘for’ and ‘of’, which have a clear implication of purposes of assessment. Since only learners can learn no matter how much feedback or guidance is provided, if the learner does not choose to use this information, there will be no learning. Because of the above all ideas of functions and/or purposes should not be and cannot be superimposed on any assignment either when it is conceived or produced or assessed. A piece of student work, as stated above, can be used for multiple functions at any stage post-production. What is important is that an assignment is assessed according to the agreed published criteria, and that these should be at a specified standard. By excluding functions and/or purposes the focus is on the relevant criteria and standards, and once the assessment has been performed the function can be locally determined. Similarly, since this research focuses on tutor understandings, the concept of ‘for’ learning being linked to formative assessment has to be classified taking into account the intentionality of the tutors, and therefore information must be translated as feedback having some impact on students.

A further question arises: if SA and FA are classified according to functions, where do criteria and standards fit into these discourses? Around the millennium, Black *et al.*, (2003) found that their belief that learners did not require an understanding of criteria to carry out self-assessment did not coincide with the reality of practice. This too can be linked directly to a lack of understanding of assessment, particularly the process of assessment. Logically people ignore what they don’t understand. Even if the criteria are not divulged explicitly, they are still necessary as implicit inside the assessors’ minds in order for any assessment to be carried out.

A good example of confusion is to be found in our data set. All our respondents use SA for end grading – this is a requirement of their job. Although 88% also use SA for mid-course grades it is seemed strange at first that 12% do not, especially given that in Bahrain, as in other jurisdictions, such as the UK, both end and mid-course grades are used for validation. The University’s regulations require the use of both the end and mid-course work for validation. For the mid-course work, it is a requirement to provide detailed feedback to students whereas for end-of-course assessments feedback is not required (in part because the assessments are not usually returned to the students). Since these practices are established, it is not possible that the 12% are unaware. It is clear therefore that they chose to define the mid-course assessments as FA rather than SA. This indicates a difference in their perceptions of SA and FA based on the function/purpose of the assessment from their point of view. The 12% may have seen the formative component of the mid-course assessment, in the form of feedback given to the students, as the defining aspect of the assessment while the other 88% found the fact that

grades were issued and taken into consideration meant that the assessment was SA. The 12% (2 respondents) actually answered ‘yes’ when asked if mid-course work was FA. They also both say that they believe FA and SA are related and that they grade FA. The question then arises, if an assessment against predefined criteria generates both a mark (or grade) that is used in validation, and feedback to students that is used by them, is the assessment SA or FA? We seem to have exposed troublesome ideas in both the respondents and in ourselves.

But this issue is much more than troublesome: different positions on theory not only change the interpretations of the above, but also influence the decisions made. If SA and FA are seen as being differentiated by their functions (Black and Wiliam 1998), then logically, in an extreme scenario, it would not be possible for SA to be used for mid-course grades if the primary function could also be seen as helping learning through FA. An extreme scenario is useful to tease out our beliefs in grey areas. Also, using FA for mid-course grading would confuse validation in jurisdictions such as that in the UK where students must be explicitly informed about the pattern of assessment before a course (module) commences. This would require duplication of work for both students and tutors if both learning and validation were required: these are the consequences of the separation of SA and FA according to functions (Taras 2009, 2012, 2016) despite claims to the contrary that within these discourses SA can and should be used to support learning (Black 2015). This dichotomy of SA and FA according to functions has created beliefs and perceptions which do not help create and sustain learning within the Assessment for Learning (AfL) discourses (Lau 2016; Taras 2009, 2016).

If a different theoretical stance is adopted where the distinction between SA and FA is based on assessment processes, then the consequences are very different. Beginning with process means that any assessment made at any point in time, according to specified criteria and agreed standards, will produce evidence that is transparent and ethical because all the elements are open to scrutiny. This summation or SA is a neutral process and results in comments and grades which explain the strengths and weaknesses of the work according to the agreed parameters (Scriven 1967; Taras 2016). The evidence may be used by learners to improve the work (FA) or to take the knowledge to benefit future work. In the latter case, and sometimes even in the former, it is may be difficult to even be aware whether or not FA has taken place. FA is an integral part of learning, but this originates in learners and not the data, and is mitigated by context and personal experience (Taras 2013, 2016).

Finally, if SA and FA are seen as complementary and inter-related processes, then both can be used at any time. More accurately, SA can be used

at any time and even if it is linked to the intention of tutors to be used for FA, it is dependent on learners for FA to be realized.

This study has shown that misconceptions concerning assessment, or diverging assessment literacies, are to be found in the Middle East, as well as in Europe (Taras 2008; Taras and Davies, 2013, 2014; Davies and Taras, 2018). Given that here we surveyed those who teach tutors about learning and teaching practice it is safe to assume that divergence exists also among tutors and students. Reaching a consensus is fraught with difficulty in part owing to entrenched views in established tutors and staff developers, supported by institutional culture. There is a parallel in medicine, which has moved in the last few decades from following localised custom-and-practice to an evidence-based approach, with outstanding results for patient outcomes. A way forward could be to target new institutions, when they arise, with the challenge of promoting a consistent assessment literacy from which dissemination could spread.

APPENDIX 1

Questionnaire on summative and formative assessment
Where 'YES – NO' or 'SURE – NOT SURE' is presented, please circle your choice.

1. Give a rough definition of summative assessment.
2. Give an example of a summative assessment task.
3. Give a rough definition of formative assessment.
4. Give an example of a formative assessment task.
5. Do you use formative assessment tasks with your students? YES – NO
6. Do you use formative assessment tasks in class? YES – NO
7. Do you use formative assessment tasks for homework? YES – NO
8. Do you keep summative and formative tasks separate? YES – NO
9. Do you conflate summative and formative tasks ('conflate' means 'place' or 'use together')? YES – NO
10. If yes, give an example.

If you use formative assessment with your students:

11. Do you tell them it will be a formative assessment? YES – NO
12. Do you explain how it will be a formative assessment? YES – NO
13. Is formative work marked? YES – NO
14. Is formative work graded? YES – NO
15. Is formative work related to summative work? YES – NO
16. If yes, how is it related?
17. Do your students carry out self-assessment? YES – NO

18. Do you present self-assessment as a formative exercise? YES – NO
19. Do you present self-assessment as a summative exercise? YES – NO
20. Does self-assessment use both summative and formative assessment? YES – NO
21. Is theory important to us as teachers? YES – NO
22. Summative assessment can be used for end of course grades. YES – NO
23. Formative assessment can be used for end of course grades. YES – NO
24. Summative assessment can be used for mid-course grades. YES – NO
25. Formative assessment can be used for mid-course grades. YES – NO
26. Summative – assesses product. YES – NO
27. Summative – assesses process. YES – NO
28. Formative – assesses product. YES – NO
29. Formative – assesses process. YES – NO
30. Summative – assesses for validation. YES – NO
31. Summative – assesses for learning. YES – NO
32. Formative – assesses for validation. YES – NO
33. Formative – assesses for learning. YES – NO
34. Summative provides useful feedback. YES – NO
35. Formative provides useful feedback. YES – NO
36. Summative and formative are different processes. YES – NO
37. Summative and formative are similar processes. YES – NO
38. I am SURE – NOT SURE how summative and formative relate to each other.
39. Students understand summative assessment. YES – NO
40. Students understand formative assessment. YES – NO
41. Students focus on summative assessment. YES – NO
42. Students focus on formative assessment. YES – NO
43. Without looking back, give a definition of summative assessment.
44. Without looking back, give a definition of formative assessment.

Thank you very much for your time and brain power.

Acknowledgements: We are very grateful to the lecturers who gave their time to participate in this research.

REFERENCES

- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher education*, 32(3), 347-364.
- Biggs, J. (1998). Assessment and Classroom Learning: a role for summative assessment? in *Assessment in Education: Principles, Policy & Practice*, 5(1), 103-110.

- Black, P. (2015). Formative assessment—an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice*, 22(1), 161-177. <http://dx.doi.org/10.1080/0969594X.2014.999643>.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Black, P., Harrison, C., & Lee, C. (2003). *Assessment for learning. Putting it into practice*. Maidenhead: Open University Press.
- Davies, M. S., & Taras, M. (2016). A comparison of assessment beliefs of science and education lecturers in a University. *Multidisciplinary Journal of Educational Research*, 6(1), 77-99. doi:10.17583/remie.2016.1766.
- Davies, M. S., & Taras, M. (2018). *Coherence and Disparity in Assessment Literacies among Higher Education Staff*. London Review of Education, 16(3), 474-490.
- Dunn, K. E., & Mulvenon, S. W. (2019). A critical review of research on formative assessments: The limited scientific evidence of the impact of formative assessments in education. *Practical assessment, research, and evaluation*, 14(1), 7. <http://pareonline.net/pdf/v14n7.pdf>.
- Lau, A. M. S. (2016). “Formative good, summative bad?” – A review of the dichotomy in assessment literature’. *Journal of Further and Higher Education* 39, 1-17.
- Price, M., Rust, C., ODonovan, B., Handley, K., & Bryant, R. (2012). *Assessment literacy: The foundation for improving student learning*. ASKe, Oxford Centre for Staff and Learning Development.
- Scriven, M. (1967). The Methodology of Evaluation. In: Tyler R, Gagne R and Scriven M (eds) *Perspectives on Curriculum Evaluation* (Chicago, R and McNally and Co), 39-83.
- Stiggins, R. (2010). Essential formative assessment competencies for teachers and school leaders. In *Handbook of formative assessment* (pp. 233-250). Routledge.
- Stiggins, R. J. (1991). Assessment literacy. *Phi delta kappan*, 72(7), 534-539.
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238.
- Taras, M. (2007). Assessment for learning: Understanding theory to improve practice. *Journal of Further and higher education*, 31(4), 363-371.
- Taras, M. (2008). Summative and formative assessment: Perceptions and realities. *Active learning in higher education*, 9(2), 172-192.
- Taras, M. (2009). Summative assessment: The missing link for formative assessment. *Journal of further and higher education*, 33(1), 57-69.
- Taras, M. (2012). Where is the Theory in Assessment for Learning? *Online Educational Research Journal*, Available at: <http://www.oerj.org> (accessed 16 December 2012).
- Taras, M. (2013). Feedback on feedback: Uncrossing wires across sectors. In *Reconceptualising feedback in higher education* (pp. 30-40). Routledge.
- Taras, M. (2015). Innovative pedagogical practices: innovations in student-centred assessment. Online. https://www.heacademy.ac.uk/system/files/maddalena_taras_final.pdf (accessed 16 November 2018).
- Taras, M. (2016). Assessment - summative and formative - some theoretical reflections. In: Malloch M, Cairns L, and O'Connor BN (eds) *Learning and Teaching in Higher Education, Volume Four: Assessment and Evaluation in Higher Education*, London: Sage, 466-478.
- Taras, M., & Davies, M.S. (2014). Perceptions and realities in assessment definitions and uses. *International Research in Education-Macrothink Institute*, 2(1), 93-102.
- Taras, M., & Davies, M.S. (2017). Assessment beliefs of higher education staff developers. *London Review of Education*, 13(1), 126-140.
- Taras, M., & Davies, M. S. (2013). Perceptions and realities in the functions and processes of assessment. *Active Learning in Higher Education*, 14(1), 51-61.
- Tight, M. (2004). Research into higher education: An a-theoretical community of practice? *Higher Education Research and Development*, 23(4), 95-411.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers’ assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40, 241-256.
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and teacher education*, 58, 149-162. <https://www.sciencedirect.com/science/article/pii/S0742051X16300907>.