Article

Corresponding Author:
Kiren Kaur d/o Ratan Singh kiren.kaur@nie.edu.sg

Formative Assessment in the English Language Classroom: A Review of Diversity and Complexities

Kiren Kaur d/o Ratan Singh
English Language & Literature Academic Group
National Institute of Education, Nanyang Technological University
Singapore

ABSTRACT

Formative assessment has been found to play a significant role in enhancing the effectiveness of classroom learning. However, it is often regarded to be a complex construct given the diversity of definitions and practices associated with it. As teachers play a critical role in the implementation of formative forms of assessment in the English language classroom, this article aims to review and explore some of the diversity and complexities involved in comprehending what formative assessment entails alongside what practices are deemed to be formative. The article argues for the need for teachers to be aware of issues and challenges associated with formative assessment and the need for them to define what they regard to be formative alongside what formative practices should look like within their own specific contexts. This is essential when teachers embark on making decisions in implementing it as part of their assessment and classroom related literacy practices.

KEYWORDS: formative assessment definitions, formative practices, English language assessment, language pedagogy

Introduction

Formative assessment (FA) is regarded to have immense pedagogical potential in reinforcing students' learning (Black & Wiliam, 1998, 2018). However, there are a variety of definitions and practices associated with it which may make it difficult for a teacher to negotiate the complex terrain of what is defined as FA and how it is to be implemented in the English Language classroom (Singh, 2019). This article aims to explore and highlight the diversity and complexities associated with FA in the English language classroom and argue for the need for teachers to define FA and

Formative Assessment in the English Language Classroom: A Review of Diversity and Complexities

its associated practices within their context of use. Awareness of issues and challenges associated with FA can possibly lead to better assessment and FA practices amongst teachers.

Origins and Development of FA

The term 'formative' was first coined by Scriven (1967) who positioned it from the perspective of programme evaluation. The terms 'formative' and 'summative' were used to describe the different roles of evaluation where formative evaluation was used for programme improvement while summative evaluation was used to judge the overall value of an educational programme (Scriven, 1967). This terminology was later applied by Bloom (1969) to educational assessments, extending it beyond the context of programme evaluation, leading to the terms 'formative' and 'summative' being employed towards different kinds of assessments (Bloom et al., 1971).

According to Bloom (1969), FA would "provide feedback and correctives at each stage in the teaching-learning process" while summative assessment (SA) aimed to judge what the student had achieved at the end of a programme or course (p.48). As the goals of both education and that of assessment have since expanded, there are now varied definitions, interpretations and practices, especially for FA. The distinctions between SA and FA have also evolved from what was first proposed by Scriven (1967), and further developed by Bloom (1969), which heightens the complexities behind the terms and especially in defining and distinguishing between the two as reflected in the section on definitions and uses of FA. The next section highlights the need for clarity in FA based on research in the field of FA practices.

Research on FA Practices

Research has indicated that it may be a struggle for FA to be enacted well if there are issues associated with teachers' uneven understanding of FA (Cotton, 2017; Deneen et al., 2019; Dixon & Williams, 2001; Kaur, 2023; Sach, 2015; Voltane & Beckett, 2011). In the context of English language learning, Dixon and Williams (2001) found that FA was understood to varying degrees in the classroom and teachers struggled to enact it suggesting that perhaps teachers were confused about the nature, purpose and effects of FA. They reported how benchmarks and interim assessments the teachers in their study used were packaged erroneously as FA, although feedback for improvement was not part of the process. They shared how a number of teachers in their study took on a summative orientation in the use of running records, a reading proficiency literacy tool that they intended to be used for formative purposes. The results were graphed as evidence however the teachers did not analyze or utilize the assessment information related to children's oral language that they obtained. The study revealed teachers' lack of feedback related practices as the results attained were merely filed. The teachers also lacked the ability to articulate or explain clearly how the tools they used assisted in language learning. It appeared that although their teachers seemed to be using a formative tool, their usage was not formative as possibly intended. Cotton (2017) similarly found that although teachers had an understanding of FA, yet they were deliberately not using data available from it to enhance student language learning.

Research additionally suggests that a lack of firm understanding of FA may mean that teachers do not see the possibility of FA being connected to SA nor would they be clear about the

distinctions between FA and SA (Brown, 2004; Gulikers, Biemans, Wesselink & Van Der Wel, 2013; Irving, Harris, & Peterson, 2011; Kaur, 2021). Brown (2004), for example, found tensions between teachers' conceptions in leveraging assessment for improvement (a more formative role) and assessment for accountability (a more summative role). This was found to be so even in low stakes exam situations in the study by Irving et al. (2011). In Kaur's (2021) study, some teachers were found to have a lack of understanding of FA and hence carried FA out in their grade 2 language classrooms much like the summative assessments they were familiar with. FA was used to churn grades for English language for record keeping purposes. Moreover, these teachers could not clearly articulate what FA meant nor how it was being used formatively. There was also an absence of the use of FA to close gaps in language learning. The confusion between FA and SA possibly arises as the purposes and intents which differentiate the two are usually not articulated clearly to teachers, which affects understanding of it (Harlen, 2005; Kaur, 2021; Singh, 2019, Yan, Zhang & Cheng, 2021). Neesom (2000) suggested the need for clear guidance in order to clarify misconceptions in distinguishing between SA and FA. The next section attempts to explain the complexities associated with FA.

Varied Definitions, Uses, Strategies and Practices Related to FA

Over the years, varied viewpoints and definitions for FA have surfaced which suggests a lack of commonly agreed upon definition and use for it (Bennett, 2011; Deneen et.al, 2019; Dunn & Mulvenon, 2009; Wininger & Norman, 2005). Wininger and Norman (2005), for example, examined over 20 commonly used textbooks and found that the significance, definition and uses of FA varied considerably and that there was no commonly agreed upon terminology in describing it.
Research also suggests that in some contexts, the term 'formative' is applied as a label for an assessment instrument or tool (Bennett, 2011). A definition of this manner is problematic considering that the same assessment instrument can potentially at times be used both formatively and summatively. Moreover, focus seems to be placed on the intention behind the assessment tool rather than how the instrument may eventually be used. Other researchers described FA as a process (Popham, 2008; Shepard, 2005; Wiliam & Leahy, 2007) and they argued that the term 'formative' should describe practitioners' uses of assessment rather than the assessments themselves. Popham (2008), for example, defined FA as a planned process in which teachers use evidence from assessments to adjust their ongoing instruction for student learning. This approach recognizes that an assessment may be formative if the teacher uses it to inform instructional choices and practices, so the manner in which an assessment is used, rather than the type of assessment itself, is what matters. In other words, as long as practitioners use an assessment to inform instructional practices, any instrument may be regarded as formative, regardless of its original intended purpose. Research similarly suggests that taking on either a process or instrumentation approach may not be ideal as process cannot rescue unsuitable instrumentation nor can instrumentation save an unsuitable process (Bennett, 2011). FA is perhaps then "best conceived as neither a test nor a process, but some thoughtful integration of process and purposefully designed methodology or instrumentation" and how the "two components work together to provide useful feedback" (Bennett: 2011: 7). These arguments, thus, impact how FA should be defined and possibly used within any particular context.

Definitions involving SA and its use on the other hand appear to be more straightforward as there is general consensus that SA serves an evaluative function with the main purpose being "measurement or judgment for accountability, certification and selection" (Swaffield, 2009: 2). According to Cizek (2010: 3), any information system is regarded as summative if it meets two conditions; one of which is that "it is administered at the end of some unit of instruction" and the other is that "its purpose is primarily to categorize the performance of a student or system". These elements are present in most definitions and uses associated with SA.

Literature also exists delineating the complexity of comparing SA and FA as the same assessment can either be used summatively or formatively (Dunn & Mulvenon, 2009; Wiliam & Thompson, 2007; Wininger, 2005). Wininger (2005), for example, used SA for formative purposes and called this 'formative summative assessment' as it provided both qualitative and quantitative feedback. Applied to the context of, for example a rubric used in a specific literacy area, this would mean that the rubric could be used summatively to summarize and measure achievement attained in that literacy area. It could also be used formatively to help develop and plan for improvement from the shortcomings surfaced, where feedback from the rubrics is used to feed forward into language instruction or improvement. So, formative and summative feedback can be provided from the same assessment tool.

The complexity in differentiating the two is moreover heightened by the fact that there is no guarantee that an assessment will be applied according to its intended purpose (Dunn & Mulvenon, 2009; Kaur & Lim-Ratnam, 2023). Dunn and Mulvenon (2009) cautioned that a test designed to give formative feedback is only formative if the teacher uses it to provide feedback for the students. A test that was intended by purpose to provide formative feedback may possibly end up being summative if no feedback is given. In some instances, even when feedback may be given, it may not support learning or it may not be used by the student to support learning. This can likely happen when general phrases like good work or well done are used by the teacher as feedback, for example, in writing and speaking tasks. In such a case, it would be difficult for such an assessment to be considered formative as no feedback to support student learning is given.

This exemplifies the complication of drawing boundaries on an assessment being either formative or summative solely based on its definition, type or use. It can, thus, be said that an assessment may be designed and packaged as FA or SA but it would be the actual methodology and use of results that determine whether an assessment is formative or summative. The categorization of FA, therefore, needs to be situated in the context of how it is enacted in the language classroom. Teachers should, therefore, consider carefully how they define and use FA in their contexts. If we are unable to clearly define FA and articulate what the process and function involving FA entails in our language classroom contexts, we may not be able to document its effectiveness meaningfully.

Operationalizing FA in the classroom alongside the use of strategies associated with FA is similarly varied and has yielded diverse classroom practices (Dixon & Williams, 2001; Gulikers et al., 2013; Sach, 2015, Yan et al., 2021). This is possibly because FA serves a variety of classroom functions (Shavelson et al., 2008). Some researchers describe the essential elements and strategies associated with FA to include the (1) identification of the learning gap, (2) feedback, (3) learning progressions, and (4) student involvement (Heritage, 2007; Sadler, 1989;

Wiliam & Thompson, 2007). Others highlight the role of questioning, feedback through grading, peer and self-assessments (Black & Wiliam, 1998; Black, Harrison, Lee, Marshall & Wiliam; 2004). Recently, Wiliam (2017), for example, suggests the following as key FA strategies that need to be in place to improve student learning:

1. Clarifying, sharing and understanding learning intentions and criteria for success.
2. Engineering effective classroom discussions, activities and learning tasks that elicit evidence of learning (developing effective classroom instructional strategies that allow for the measurement of success)
3. Providing feedback that moves learning forward.
4. Activating learners as instructional resources for one another (students working with each other in discussions and in groups).
5. Activating learners as owners of their own learning.

Furthermore, in attempting to operationalize or use FA strategies in the classroom, Shavelson et al. (2008) suggested that teacher practices associated with FA could be viewed from the perspective of phases of instruction. They highlighted how FA practices could fall on a continuum depending "on the amount of planning involved, its formality, the nature and quality of the data sought, and the nature of the feedback given to students by the teacher" (Shavelson et al., 2008: 300). Three types of FA practices are described, that being "on-the-fly", "planned-for-interaction", and "embedded-in-the-curriculum" assessments (p.300). Figure 1 below reflects their continuum based on these three assessment types.

Figure 1: Variation in FA Practices (Shavelson et al., 2008: 300)

Shavelson et al. (2008: 300) explained that "on-the-fly" FA is informal and unplanned as it occurs spontaneously when there is a "teachable moment" in the classroom. FA practices in the middle of the continuum fall under "planned-for-interaction" as these are more deliberate in comparison and involve questioning designed to discern and improve students' knowledge acquisition (p.300). In such assessments, the teacher "plans for and crafts ways to find the gap between what students know and what they need to know" (p.300). At the other end of the continuum are "embedded-in-the-curriculum" FA practices and these are "ready-to-use" and are "embedded at junctures" within a unit where it is important that a sub-goal is reached before students go on to the next lesson (p.301). Embedded assessments are planned and help tell the teacher "about what students currently know, and what they still need to learn (i.e., 'the gap') so that teachers can provide timely feedback" (p.301). Regardless of which FA approaches and strategies are advocated in the language classroom, in taking the perspective of the role of the teacher in FA (or what a teacher does with FA in the classroom), teachers need to note that there is in general a consensus on the importance of feedback provided in order to move learning forward or to close learning gaps (Singh, 2019).

There are, thus, diverse definitions, uses, strategies and practices a teacher needs to consider in implementing FA in a given classroom context. FA is then, perhaps, best seen as teachers being clear about how they perceive FA, what they would like FA to do alongside how they plan to use FA strategies and approaches in the language classroom. The intent of the article is not to press forth one particular view of FA, what it is and what it should be in various contexts of use or across the various literacy areas. The article aims to create an awareness amongst teachers of the complexities associated with FA and the need to define it (while taking into account their needs and intents in using FA) based on what they deem to be appropriate for their specific language or literacy contexts. Thinking about FA can possibly be considered based Figure 2 below which attempts facilitate the process of mapping out key considerations to take in FA enactment. In terms of the approaches and phases of instruction planned out, possibly "on-the-fly" FA can be left out due to its informal and unplanned nature and it can be assumed that it occurs spontaneously when a teachable moment arises in the language classroom (Shavelson et al., 2008).

Definition of FA	Intent of FA	Approaches, Phases of Instruction and FA Strategies Used
How do you define FA in your English Language Classroom?	What is the intent of using FA in the language classroom?	1) Which literacy area are you planning to use FA for? 2) Is the focus based on a 'Planned-for-Interaction'? Is the focus on an 'Embedded Assessment'? 3) What strategies and practices are deemed to be appropriate based on how you intend to use it in your context of use based on phases of instruction, which is appropriate? 4) How do you intend to use it in this literacy area or in your lesson? 5) Is the feedback attained utilized for student learning or for closing gaps in learning?

Figure 2: Planning for the use of FA in the language classroom

Figure 2 above, therefore, attempts to articulate key aspects in need of consideration when using FA in the classroom so that teachers are better able to make informed decisions when implementing it as part of their assessment and classroom related literacy practices.

Recommendations

As the article argues for the need for teachers to be aware of issues and challenges associated with formative assessment and the need for them to define what they regard to be formative alongside what formative practices should look like within their own specific contexts, three recommendations are made towards implementing FA in the language classroom.

One would be the need for teachers to establish a clear lexicon of FA, its definition, corresponding strategies alongside what they deem to be FA practices based on the teachers' intents and needs within their school and language classroom contexts (Singh, 2019). A clear definition could be articulated across the school context while some strategies or practices could be planned out for language teaching areas or across various grade levels. These could vary based on the levels the teachers teach, the literacy areas being focused on and even across different classroom contexts within a particular grade level. Figure 2 above could be used to facilitate planning for FA in the language classroom or for a particular grade level or literacy area. Samples of what FA could look like alongside how these can be used in the classroom could be discussed, articulated and shared. Scoping FA based on what is relevant for the context of use is important given the diversity associated with FA.

The other would be the need for more sharing and discussions to be carried out on FA and how it is implemented in various language classroom contexts across national or international platforms so that teachers can learn from other teachers and contexts as a community. Sharing sessions could encompass teachers imparting ideas on their best practices associated with FA or how they have carried out FA within literacy areas and language grades in their contexts based on how they define FA. Such sessions will provide valuable insights to others wanting to implement FA as well as create avenues for professional learning and development as a community of teachers.

The third recommendation would be the need to ensure that teachers have adequate assessment literacy in general in terms of the diversity and complexities of FA, what FA entails and what it looks like within specific language classroom contexts (Kaur & Lim-Ratnam, 2023). Assessment literacy on FA needs to be raised and these could be through various professional development avenues available to the teachers. The teachers could also create professional learning communities within their school contexts to share and learn together as a community on how FA can be carried out within their language or specific literacy area contexts based on the grade levels they teach. Knowledge and sharing knowledge are empowerment and this could lead to teachers having a better understanding of what FA encompasses, how to define it and how best to use it in their contexts. Without adequate subject matter knowledge, teachers will not be able to define FA for their context of use or develop a repertoire of FA strategies for the various literacy areas for student learning purposes.

Conclusion

In acknowledging the widespread recognition of the significance of FA on classroom learning, this article does not aim to impose a certain position on FA, what it should or should not be. It aims to interrogate and highlight the diversity and complexities associated with FA to create awareness of the complex terrain teachers negotiate as they embark on making FA decisions when implementing literacy practices in their classrooms. The article aims to enhance a teacher's ability in making better informed decisions on how to implement FA as part of their assessment and classroom related literacy practices. As FA is an important area of research, future directions can include studies on how English teachers define FA within a given context and what they deem to be their

associated FA practices within it. This will help to further interrogate the issue of how definitions translate to teacher practices.

References

Bennett, R.E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy and Practice, 18*(1), 5-25. https://doi.org/10.1080/0969594X.2010.513678

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1)*,* 7-74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 1-25. https://doi.org/10.1080/0969594X.2018.1441807

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi delta kappan*, *86*(1), 8-21.

Bloom, B.S. (1969). Some theoretical issues relating to educational evaluation. In R.W. Tyler (Ed.), *Educational Evaluation: New Roles, New Means. The 63ʳᵈ Yearbook of the National Society for the Study of Education, part 2, 69*, (pp. 26-50). Chicago: University of Chicago Press.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook of Summative and Formative Evaluation of Student Learning*. New York: McGraw-Hill

Brown, G. T. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, *11*(3), 301-318. https://doi.org/0.1080/0969594042000304609

Cizek, G. J. (2010). An introduction to formative assessment: History, characteristics, and challenges. In H., Andrade & G.J. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 3-17). New York: Routledge.

Cotton, D. (2017). Teachers' use of formative assessment. *Delta Kappa Gamma Bulletin*, *83*(3).

Deneen, C.C., Fulmer, G.W., Brown, G.T.L., Tan, K., Leong, W.S. & Tay, H.Y. (2019). Value, practice and proficiency: Teachers' complex relationship with assessment for learning. *Teaching and Teacher Education: An International Journal of Research and Studies*, 80(1), 39-47. Elsevier Ltd. Retrieved on September 26, 2022 from https://www.learntechlib.org/p/202097/.

Dixon, H., & Williams, R. (2001, September). *Teachers' understandings of formative assessment.* Paper presented at the Annual Conference of the British Educational Research Association. Retrieved from http://www.leeds.ac.uk/educol/documents/ 00002533.htm

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research& Evaluation, 14*(7), 1-11.

Gulikers, J. T., Biemans, H. J., Wesselink, R., & Van Der Wel, M. (2013). Aligning formative and summative assessments: A collaborative action research challenging teacher conceptions. *Studies in Educational Evaluation*, *39*(2), 116-124. https://doi.org/10.1016/j.stueduc.2013.03.001

Harlen, W. (2005). Teachers' summative practices and assessment for learning–tensions and synergies. *Curriculum Journal*, *16*(2), 207-223. https://doi.org/10.1080/09585170500136093

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, *89*(2), 140-145. https://doi.org/10.1177/003172170708900210

Irving, S. E., Harris, L. R., & Peterson, E. R. (2011). 'One assessment doesn't serve all the purposes', or does it? New Zealand teachers describe assessment and feedback. *Asia Pacific Education Review*, *12*(3), 413-426. https://doi.org/10.1007/s12564-011-9145-1

Kaur, K. (2021). Formative assessment in English language teaching: exploring the enactment practices of teachers within three primary schools in Singapore. *Asia Pacific Journal of Education, 41*(4), 695-710. https://doi:10.1080/02188791.2021.1997707

Kaur, K., & Lim-Ratnam, C. (2023). Implementation of formative assessment in the English language classroom: insights from three primary schools in Singapore. *Educational Research for Policy and Practice*, 22, 215-237. https://doi.org/10.1007/s10671-022-09327-y

Neesom, A. (2000). *Report on Teachers' Perception of Formative Assessment.* London: Qualification and Curriculum Authority.

Popham, W.J. (2008). *Transformative Assessment.* Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

Sach, E. (2015). An exploration of teachers' narratives: what are the facilitators and constraints, which promote or inhibit 'good' formative assessment practices in schools? *Education, 3-13*, *43*(3), 322-335, https://doi:10.1080/03004279.2013.813956

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119-144.

Scriven, M. (1967). The methodology of evaluation. In R.W. Tyler, R.M. Gagne, & M. Scriven (Eds.), *Perspectives of Curriculum Evaluation* (pp. 39-83). Chicago: Rand McNally.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A. Tomita, M.K., & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, *21*(4), 295-314. https://doi.org/10.1080/08957340802347647

Shepard, L.A. (2005, October). *Formative assessment: Caveat emptor.* Paper presented at the Educational Testing Service (ETS) Invitational Conference 2005: The Future of Assessment: Shaping Teaching and Learning. New York. Retrieved from http://www.cpre.org/ccii/images/stories/ccii_pdfs/shepard%20 formative %20assessment%20caveat%20emptor.pdf

Singh, K. K. R. (2019). *Enactment of formative assessment in the lower primary English classroom: case studies of teachers' practices in three schools* (Doctoral dissertation). Retrieved from: http://hdl.handle.net/10497/21223

Swaffield, S. (2009, September). *The misrepresentation of Assessment for Learning and the woeful waste of a wonderful opportunity.* Paper presented as a 'work in progress paper', at the Association for Achievement and Improvement through Assessment (AAIA) National Conference. United Kingdom: Bournemouth. Retrieved from: http://www.aaia.org.uk/content/ uploads /2010/07/The-Misrepresentation-of-Assessment-for-Learning2.pdf

Voltane, L. & Beckett, D. (2011) Formative assessment and the contemporary classroom: Synergies and tensions between research and practice. *Canadian Journal of Education*, *34*(2), 239-255. Retrieved from: https://www.jstor.org/stable/10.2307/canajeducrevucan.34.2.239

Wiliam, D., & Leahy, S. (2007). A theoretical foundation for formative assessment. In J. H. McMillan (Ed.), *Formative Assessment Classroom: Theory into Practice* (pp. 29–42). New York, NY: Teachers College Press.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it take to make it work? In C.A. Dwyer (Ed.), *The Future of Assessment: Shaping Teaching and Learning* (pp. 53-82). Mahwah New Jersey: Lawrence Erlbaum Associates.

Wininger, S. (2005). Using your tests to teach: Formative summative assessment. *Teaching of Psychology*, *32*(3), 164-166. https://doi.org/10.1207/s15328023top3203_7

Wininger, S., & Norman, A. (2005). Teacher candidates' exposure to formative assessment in educational psychology textbooks: A content analysis. *Educational Assessment, 10*(1),19-37. https://doi.org/10.1207/s15326977ea1001_2

Wiliam, Dylan (2017). *Embedded Formative Assessment: (Strategies for Classroom Formative Assessment that Drives Student Engagement and Learning)*. (2nd Ed). New Art and Science of Teaching.

Yan Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice, 28*(3), 228-260. https://doi.org/10.1080/0969594X.2021.1884042

Yan, Q., Zhang, L. J., & Cheng, X. (2021). Implementing classroom-based assessment for young EFL learners in the Chinese context: a case study. *The Asia-Pacific Education Researcher*, *30*, 541-552. https://doi.org/10.1007/s40299-021-00602-9