OPEN ACCESS

UNIVERSITY OF BRISTOL

Gu, P. Y., & Yu, G. (2020). Researching classroom-based assessment for formative purposes. *Chinese Journal of Applied Linguistics*, *43*(2), 150-168. https://doi.org/10.1515/CJAL-2020-0010

Peer reviewed version

Link to published version (if available):
10.1515/CJAL-2020-0010

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Researching Classroom-Based Assessment for Formative Purposes

Peter Yongqi Gu
Victoria University of Wellington, New Zealand

Guoxing Yu
Bristol University, United Kingdom

Abstract

This article aims to help teacher-researchers engage in empirical research on classroom-based assessment for formative purposes. We will first introduce the key features of classroom-based formative assessment (CBFA), and analyze the research questions asked in round-one projects funded by the Fund for Assessment Research (FAR) in Foreign Language Education in China. Next, we will illustrate how some research questions can be answered by analyzing a video-taped lesson from a round-one FAR project. We conclude by calling for more teacher-led research and argue that research on CBFA by teachers will not only produce valid interpretations and applicable findings, but also constitute a viable model for teacher professional development.

Keywords: formative assessment, classroom-based assessment, teacher research

1 Introduction

Over the last two decades since the publication of Black and Wiliam's (1998) highly influential article, formative assessment has become a popular topic in educational reform. Teachers are being asked to implement formative assessment in their classrooms. They are also encouraged to conduct their own research on formative assessment. A major problem that has been encountered in both implementation and research is an operationalizable definition of formative assessment. A related problem is the unit of analysis for classroom assessment research. In working with teachers and teacher researchers over the last few years for the implementation of various assessment research agendas, we have repeatedly felt the urgent need for teacher literacy in researching classroom-based formative assessment (CBFA). This article arises from this practical need. We will provide an operationalization framework and offer insights into the processes in conducting CBFA research.

Assessment can become formative when evidence of learning is elicited and matched against the learning target to inform the learner about the gap between the learner's current state of knowledge or ability and the target. To be really helpful in closing the gap, a formative assessment event needs to be rounded off with action. Davison and Leung (2009) outline two basic functions of formative assessment: informing and forming. The former

puts emphasis on the necessary but insufficient nature of feedback, while the latter underscores the importance of follow-up action in order for learning to take place.

Similarly, Andrade (2010) simply conceptualizes formative assessment as "informed action" (p. 345). Expressed in another way, most researchers (Black & Wiliam, 2012; Ramaprasad, 1983; Sadler, 1989) believe that the essence of formative assessment involves establishing 1) where the learners are going; 2) where the learners currently are in their learning; and 3) what needs to be done to get them there.

While teacher research is being encouraged as part of an international effort in capitalizing on the powers of formative assessment, the contingent nature of CBFA makes it very hard for teachers to study formative assessment in their own classrooms. This article attempts to help those teachers who want to do empirical research on formative assessment that happens in classrooms. In doing so, we first of all delimit the parameters of CBFA. Next, in order to illustrate the *what* and the *how* of research questions on CBFA, we analyze and critique the research questions asked in a set of research projects led by teachers at both secondary and tertiary levels in China. Finally, we illustrate one way in which CBFA can be researched by analyzing a video-taped lesson at a secondary school in North China.

2   Classroom-based formative assessment

CBFA is a teaching/learning event that happens within or beyond one class. The event includes 1) elicitation of evidence of students' understanding or learning, 2) interpretation of the elicited information against the learning target and success criteria, 3) feedback based on this interpretation for the student in question, and 4) follow-up action taken by the student or teacher to improve learning. All these elements must be present before each CBFA event is complete. And more often than not, learning takes place after a series of these cyclical and spiraling CBFA events.

Classroom assessment practices that involve elicitation of evidence, interpreting the evidence, providing feedback, and student/teacher take-up and action form one complete CBFA event. Each event is aimed at one target of learning, teaching, and assessment; and each step or element has the learning target as the reference point. These elements are both sequential and interactive. The completion of one cycle normally will necessitate a readjustment of the target which entails another cycle of assessment practice. The elements, therefore, form spiraling cycles, with each complete cycle moving student understanding or learning closer to the target. This happens continuously until a judgment is made that the target is reached and the success criteria met.

Depending on the scope of the task being assessed, a complete cycle of an assessment event mentioned above can take a few seconds; or it may take a week or much longer to complete. Wiliam (2010) groups the lengths of these cycles into three types: short-, medium-, and long-cycles. Short cycles happen minute-by-minute and day-by-day; medium cycles are completed within one to four weeks; and long cycles can take anywhere between a month to a year to be completed.

CBFA normally belongs to the "short-cycle" category. This is especially true for those assessments that happen within the classroom. That said, learning usually takes place in

timespans longer than a normal class. It is, therefore, often the case that teachers and learners need to check again and again in order to see the effect of learning and see if a course of action works. These actions would take longer than one class and can also be regarded as CBFA. Formative assessment events that go beyond a month or so to complete are normally more formal. For example, information from a formal diagnostic test can be used to guide learning efforts for a whole semester or more. These normally happen well beyond regular classes, and, despite being formative in nature, cannot be counted as CBFA anymore, simply because most of the assessment practices do not happen inside the classroom.

Inside the classroom, many assessment opportunities arise spontaneously without the teacher's preparation. These normally take the form of classroom interactions or the teacher's observations of the students' task performances. Cowie and Bell (1999) labeled these assessment events "interactive." Interactive formative assessment events are usually triggered by the teacher noticing an unexpected or erroneous understanding or performance. On the spot interpretation of the deviant understanding would help the teacher recognize the error as a significant point to focus on. The teacher may immediately ask another student the same question and see if the problem is pervasive (both a follow-up action of the previous assessment event and the start of another assessment event). If the gravity of the problem is deemed serious, the teacher may decide to explain, re-teach, or change a practice activity for the whole class.

The same phenomenon has been observed by Ruiz-Primo and her colleagues who labeled it "informal formative assessment" (Ruiz-Primo, 2011; Ruiz-Primo & Furtak, 2006, 2007). These researchers developed this into an observation framework that included eliciting (E), student response (S), recognizing (R), and using information (U) and called it the "ESRU cycle." Interestingly, their studies indicated that informal teacher classroom assessment practices include different configurations in terms of how many elements are practiced. Few complete cycles of informal formative assessment were found. Instead, teachers used ES more often than ESR and ESRU. Those who used more complete ESRU cycles were found to benefit their students better.

An overwhelming proportion of assessment activities happening in classrooms are contingent, and the cycles are short and often incomplete. The formal, semi-formal, and often curriculum-embedded assessment activities in or out of everyday classes can be used for formative purposes as well.

3   Researching classroom-based formative assessment

When engaging in research projects on classroom-based formative assessment, we often hear teachers asking the following questions:

- Is what I'm doing research?
- What does research look like?
- How do I do classroom-based research on formative assessment?

In this section, we begin by briefly outlining the common types of research related to CBFA, followed by an analysis of research questions. We will end the section with an example of

one video-taped lesson, and illustrate what research questions can be asked about this lesson and how these exploratory research questions can be answered.

3.1   Common types of research

It might be reassuring to note that we have only a limited repertoire in terms of the kinds of research questions to ask, the methods we use to collect data, and the ways we analyze the data in order to answer the research questions. Table 1 presents a rough classification of the types of research that can be done on CBFA. Types of research questions are listed in the left column, followed by the potential kinds of data that can be collected to answer each type of research question, plus potential data analysis that can be done.

When we want to know what assessment practices can be found in a teacher's classroom, we need open-ended explorations such as classroom observations. This kind of research is called *exploratory* research. In exploring a teacher's classroom practices, we often need to describe in detail the kind of classroom discourse, teaching procedures and artifacts for teaching. From this perspective, we call this *descriptive* research. In answering exploratory and descriptive research questions, we often use data collection tools such as audio/video recordings and observation sheets. Another common tool is interviews of teachers and learners about their perceptions and beliefs. Once we have the data at hand, we normally code the recordings and interviews for analysis and interpretation. After the coding, codes can be tallied and analyzed quantitatively if there are a substantial quantity of cases. Alternatively, the data can be presented qualitatively and narrated or explained through a particular theoretical framework.

Explorations and descriptions of what we want to know normally fall into the beginning stage of understanding. We often want to know more about the issue, such as the nature of the relationship among various factors involved. This type of research questions are *relationship* questions. For CBFA, for example, we might want to know which factors are related to the teacher's successful implementation of formative assessment and which are not, or whether a teacher's use of feedback is related to subsequent learning behaviors. To establish a relationship, large-scale data would be suitable for charting patterns. We can use a questionnaire to collect the data, and perform a correlation analysis to discover how closely the factors are associated with each other.

Correlational research described above does not tell us if a cause and effect relationship exists between two things. For example, if we obtain a strong correlation between a teacher's formative assessment practices (e.g., types and frequencies of feedback) and the students' exam scores, we can only infer from the correlation that the two are closely related, but are not sure if the student scores are a result of teaching practices. To answer research questions that are cause and effect in nature, we can either design an educational experiment where we try to single out the effect of our innovation by building into the design a pre-test and a post-test, so that we can see whether there is concrete evidence of improvement. We also need a comparison or control group who are comparable to our experimental group in every way but do not receive our innovative treatment. The control group ensures that the improvement of our students' scores at the end of the experiment has indeed been the result of our classroom innovation and not due to, for example, pre-existing differences.

Table 1. *Researching classroom assessment: What and how*

What			How: collecting data	How: analyzing data

Descriptive and exploratory			• Classroom audio/video recordings
			
			• Observation sheets
			
			• Teacher/Learner interviews
			
			• Journal logs
			
			• Lesson plans, exercise sheets, and other artifacts
			
			• Questionnaires (if large scale)
				• Coding and analysis of
				
				  o discourse structures, e.g., IRF
				
				  o activity types
				
				  o elements of formative assessment
				
				• Descriptive statistics

Relationship	Correlational		Surveys Tests	Correlations Multiple regressions
	Cause and effect	Experimental	intervention, pre- and post-tests, plus control group comparison	Inferential statistics: e.g., t-test, ANOVA, ANCOVA, MANOVA
		Action research	Innovation→observation→evaluation→ CHANGE→observation→evaluation→CHANGE→ observation→evaluation…	• Thick description and narration • Statistical comparison (if possible)
Comparison			All comparable data types	Qualitative comparisons and Quantitative comparisons (t-test, ANOVA, etc.) between different groups (male vs. female, high vs.

low proficiency; urban vs. rural
schools; etc.) and factors.

For teachers trying to implement CBFA, perhaps the most often asked research question would be "Will CBFA be useful for my students?" To answer this question, arguably the most suitable research method for the classroom teacher is *action research*. The teacher's action research typically involves a cyclical process whereby the teacher systematically and continually monitors his or her classroom innovations and uses the feedback to improve teaching. These spiraling cycles appear in the following sequence: innovation (CBFA), observation, evaluation, CHANGE, observation, evaluation, CHANGE, observation, evaluation. At the end of the first cycle, decisions are made as to what new activities and procedures and refinement of existing practices should be carried out next, based on the evaluation of the effectiveness of CBFA activities in the first round. This triggers the start of a revised cycle of CBFA. The spiraling cycles go on until the teacher is satisfied with the evidence obtained showing the improvement of teaching and learning. In action research, the data collected can be anything the teacher deems relevant; and the analysis tends to be more qualitative, descriptive, or narrative than quantitative.

Another common type of research relevant to CBFA is *comparison research*. This type of research aims to discover similarities and differences between different groups and factors. Questions such as the following are typically asked: Do male and female students differ in their perceptions of and reactions to CBFA? Do high proficiency students in the class benefit from CBFA more than low proficiency students? To what extent do urban and rural schools differ in their receptiveness to CBFA? All useful data types can help answer these questions. Thick descriptions and narrations can be used to analyze qualitative data. Statistical comparisons (e.g., t-test, ANOVA) can be employed for the comparison of quantitative data.

3.2   Asking research questions

A crucial first step in our research as we conceptualize a project is the formulation of research questions. Research questions operationalize the research problem/issue. As such, they are concrete, focused, and empirically answerable. One way to ensure a good research question is to list the research methods alongside this research question.

3.2.1   Research questions in FAR-funded projects

To illustrate the common problems in asking research questions, we analyzed the research questions in eight research projects funded by the Fund for Assessment Research in Foreign Language Education. In addition to what we normally call "research questions," we found three other types of questions listed as research questions: teaching questions, literature review questions, and other questions that we labeled "extension questions." Table 2 illustrates these types of questions with fictitious examples.

Table 2. *Distinguishing research questions from other questions*

Types of questions	Research questions	Teaching questions	Literature review questions	Extension questions
Example	Does the keyword method lead to better vocabulary retention than the semantic map method?	How do I teach vocabulary?	What do the experts say about vocabulary learning?	What are the pedagogical implications of my study?

An important distinction between a research question and other questions is its empirical answerability, that is, whether we can collect data to answer it. If we look at the following research question from one of the eight studies mentioned above, we can see that, as it stands, it is not directly answerable:

How can diagnostic results better serve the teaching of reading?

The question essentially asks about how to teach reading. Potential answers to this question comes from the teacher's thinking, preparation and action; and there is no way to prove that these ways of teaching will "better serve" the teaching of reading. In order to turn this question into a researchable question that can be answered with data, we can list the potential data and the analysis of the data alongside the alternative questions (Table 3).

Table 3. *Turning a teaching question into researchable questions*

Potential new research questions	Potential data	Possible analysis
What are the diagnostic test results?	Diagnostic test results	• Descriptive statistics outlining percentages, mean scores, strong and weak areas, etc. • Qualitative interpretations and descriptions of diagnostic results
To what extent can diagnosis-based teaching of reading improve students' reading grades?	• Intervention (e.g., feedback based on diagnosis; feedback + follow-up exercises) • Pre- and post-reading tests • Student interviews	• Statistical comparisons of pre- and post-tests and of different intervention methods. • Action research and thick descriptions of qualitative data

In addition to asking unanswerable questions, another common problem is vague and broad research questions. The following research question seems largely answerable. However, many variables are combined into one question, making it hard to answer.

Can feedback provided by the diagnostic assessment tool UDig improve senior secondary school students' learner autonomy and their sub-skills in writing (for example, verb collocations, vocabulary and grammar, cohesion and coherence at the discourse level, etc.) so as to improve their writing proficiency in English?

We do know that this study is interested in finding the benefits of feedback provided by the UDig platform; and this means that the study is largely a relationship and cause and effect

study. If we analyze this research question, we see one set of independent variables, UDig feedback, two sets of dependent variables, i.e., autonomous learning ability and English language writing proficiency, plus one set of intervening variables, sub-skills in writing. Based on the information above, we will try to reformulate the research question into the following research questions (Table 4).

Table 4. *More concrete and focused research questions*

Potential new research questions	Potential data	Possible analysis
Can feedback provided by the diagnostic assessment tool UDig improve senior secondary school students' learner autonomy?	• Feedback data from UDig • Pre- and post-student interviews or autonomous learning questionnaire	• Action research • Matching types of feedback with types of autonomous learning behavior
Can feedback provided by the diagnostic assessment tool UDig improve senior secondary school students' writing proficiency in English?	• Feedback data from UDig • Pre- and post- writing proficiency measures or process portfolio of writing samples	• Action research • Feedback types matched with proficiency groups • Descriptions of writing improvement linked to feedback
Can feedback provided by the diagnostic assessment tool UDig improve senior secondary school students' sub-skills in writing?	• Feedback data from UDig about sub-skills • Measures of each sub-skill at multiple times	• Action research • Track feedback adjustments and link to improvement in each sub-skill
Does improvement of sub-skills, if any (e.g., verb collocations, vocabulary and grammar, cohesion and coherence) lead to the improvement of overall proficiency in writing?	• Measures of each sub-skill at multiple times • Writing proficiency measures at different times, or • Process portfolio of writing samples	• Correlation between each sub-skill and overall writing proficiency • Analysis of covariance • Thick descriptions of improvement

As Table 4 shows, turning the original broad research question into the four concrete research questions enables the researcher to plan clearly what data and its analysis best suit the purpose of the study.

3.2.2 Research questions about classroom-based formative assessment

In studying CBFA, many research questions can be asked. The following table (Table 5) provides a number of potential questions that can be asked. The list is not exhaustive, of course.

Table 5. *Questions to ask about the classroom formative assessment cycle*

Steps in an assessment event	Teacher	Students
Having clear learning target/success criteria	• Is the teacher clear about long-term, mid-term, short-term goals for teaching? • How does the teacher make success criteria clear to the students?	• Are the learners clear about their long-term, mid-term, and short-term goals? • Are the learners clear about the success criteria for the task they are performing?
Elicitation	How does the teacher elicit evidence of student learning?	What do the students do to show their current level of learning?
Evaluation and interpretation	How does the teacher interpret the evidence of learning elicited?	Are the learners engaged in self-assessment and peer-assessment?
Feedback	How does the teacher provide feedback?	How is feedback received by the learners?
Follow-up action	• Is there follow-up action by the teacher? • What is done after feedback is provided?	What actions do the learners engage in to act on the feedback received?

Since CBFA is seen as spiraling cycles of classroom assessment events, one way of studying CBFA is to examine each component or step in an assessment event. Exploratory and descriptive research questions about these components will help understand if CBFA is present, and how it works or does not work in fulfilling the assessment function.

To a large extent, research questions very much define the research methods that can be used. In other words, the most suitable research methods are chosen based on the research questions we are asking. If we focus on teacher beliefs, interviews and surveys might be fine. If we focus on assessment practices inside the classroom, direct observations plus video/audio recordings should give us insights. The following section illustrates a way in which classroom-based assessment practices can be studied.

3.3 Answering research questions: An example

Once data are collected, there are many ways we can analyze the data in order to answer the research questions. For classroom-based assessment events, we can analyze the purpose, procedures, intended effect and perceived effect (e.g., Torrance & Pryor, 2001). If we need to see a classroom discourse perspective, we can analyze the questioning (Heritage & Heritage, 2013) and Initiation-Response-Feedback (IRF) patterns (Sinclair & Coulthard, 1975). This section briefly shows a way analysis of CBFA can be done for the analysis of classroom recordings. Due to the illustrative nature of this section, answers to the research questions and the entailing discussion will be deliberately brief.

3.3.1 The lesson

A 40-minute, Year 2 senior secondary lesson was video-recorded. The school was a star school in a northern Chinese city. The textbook was *New Senior English For China* published by People's Education Press; and the lesson, "A letter of advice," formed part of Unit 3, Optional Volume 6. The lesson was part of a FAR-funded project focusing on the role of diagnostic assessment in vocabulary teaching. The focus of the lesson was to teach vocabulary through writing. It was designed as a part of a larger follow-up package after finding vocabulary weaknesses from an initial round of diagnostic tests on the FAR platform. In particular, one major weakness found in the diagnostic tests was the students' weaknesses in making use of vocabulary chunks in their own compositions. The present lesson was then designed to teach the following four expressions selected from the text: "Decide on…," "Every time…," "Instead of…," and "If you feel…."

The lesson was transcribed for analysis. Table 6 is an excerpt that shows how the transcription was done.

Table 6. *Excerpt of lesson transcript* (S=1 student; SS: two or more students; T=teacher)

Turn	Timespan	Content
1	0:06.4 - 0:15.2	S: Stand up
2	0:15.1 - 0:18.0	T: Good morning everyone
3	0:18.0 - 0:19.8	SS: Good morning professor
4	0:19.8 - 0:34.9	T: (inaudible) As you know, everyone wants to live a healthy life. But, why? Because without a healthy body, nothing is possible. Yes?
5	0:34.8 - 0:36.2	SS: Yes
6	0:36.2 - 0:52.7	T: But how? Can you give me some suggestions? There is no need to put up your hand, just stand up
7	0:52.7 - 1:09.2	S: Ok. First you need to have a balanced diet; and live a regular life
8	1:09.2 - 1:10.8	T: Next step
9	1:10.8 - 1:13.6	S: You should have a positive attitude towards life
10	1:13.6 - 1:17.2	T: Yes
11	1:17.2 - 1:24.5	T: You are the best. Anybody else?
12	1:24.5 - 1:47.2	S: We need to sometimes […] (inaudible). Enjoy the happiness and challenge in our everyday life
13	1:47.2 - 1:51.5	T: That's right. Next one?
14	1:51.4 - 2:03.5	S: Exercise

| 15 | 2:03.4 - 2:18.6 | T: Do exercise. Very good. Now we can say you'd better be positive, right? and (...) good habit, or do exercise, work out every day. Now my question is: do you take exercise every day? |

The lesson can be broken up into the following chunks: warming up and contextualization; reading Li Hua's letter pleading for advice; preparing for a reply letter; writing the reply letter; and assessing the reply letter.

3.3.2   Research questions

Our main purpose is to explore the classroom-based assessment practices in this lesson. This purpose can be operationalized into the following research questions:

1) How did the teacher elicit student understanding and learning?
2) How did the teacher interpret the students' understanding and learning?
3) What types of feedback were provided?
4) What follow-up actions were taken after the feedback?

NVivo 12 was used to transcribe and code the data. The coding system was derived from both top-down and bottom-up processes. The four-way breakdown (elicitation, interpretation, feedback, and action) of classroom assessment practices was the starting point that guided the top-down perspective of what was involved in CBFA. Exactly what the teacher did in class in each of the four components was analyzed turn by turn in order to derive the coding system. After the coding was done, a simple tally of each code was used as an indicator of the teacher's classroom assessment practices in this lesson.

It should be noted that this is an illustrative example only, and that only one lesson taught by one teacher was coded for one round by one researcher. To be able to claim representativeness to a certain extent, more lessons representing this teacher's assessment practices in teaching different types of lessons, and data from other teachers representing the target population should be included. In addition, the coding system derived from this lesson will necessarily be insufficient, and needs to be revised when other lessons from both the same teacher and from other teachers are coded. At least a second coder needs to be involved to gain inter-coder reliability and minimize subjectivity in coding.

3.3.3   Findings

Assessment practices

Both planned and contingent (or interactive) assessment practices (Cowie & Bell, 1999) can be found in this lesson, although the overwhelming majority of assessment practices fell under the contingent category (Table 7). Planned assessment happened mainly at the end of the lesson, Turn 141 onwards, when the teacher introduced criteria for assessing the letter the students had just finished writing. She then asked the students to use the criteria to do self-assessment and peer-assessment. The lesson was rounded off with class- and teacher-assessment of selected "best pieces" from various groups.

When each instance of elicitation, interpretation, feedback, and action was counted, the following figures were obtained:

- Elicitation   115

- Interpretation 77
- Feedback   57
- Action   1

Overall, we can see that a lot of elicitation was done. Not every elicitation was interpreted, and not everything elicited received feedback. If the lesson as a whole is seen as a planned follow-up lesson in order to tackle the vocabulary chunk problems found in the previous UDig tests, the whole lesson is the Action. In this sense, the lesson itself can be counted as part of the medium- or long-cycle type of formative assessment (Wiliam, 2010). This pattern of classroom assessment practices is very much in line with Ruiz-Primo and Furtak (2007) who found very few complete CBFA cycles.

Like in most other cases, contingent, interactive classroom assessment accounted for most of the assessment practices in this lesson. Table 7 outlines each class activity, the major discourse pattern for each activity, and three components of formative assessment for each stage of the lesson.

Table 7. *Elements of formative assessment in the lesson*

Lesson structure				Classroom formative assessment		
Turn	Time	Activity	Discourse pattern	Elicitation of understanding or learning	Feedback	Follow-up
4-18	2 min	Warm up: Ask students to brainstorm on a healthy life style	T→SS, T→S, SS→T, S→T	Questioning	**Confirm**: "Yes," "Ok," "That's right" **Praise**: "You are the best"	Nil
19-22	1 min	Read Li Hua's letter pleading for help	T→SS, T→S, SS→T, S→T	• Read aloud • Questioning ("Is it clear?")	**Praise**: "Your tone and pronunciation is amazing"	Nil
23-54	5 min	Prepare for a reply letter 1: Find do's and don'ts in the textbook on p.19	T→SS, T→S, SS→T, S→T	Find sentence in textbook and read aloud to class	**Confirm**: "Yes," repeating words from students **Praise**: "Good," "Wonderful" **Awareness-raising**: writes on bb; "These are very useful suggestions, and when you give suggestions, these can make your writing very coherent, very smooth like a ...."	Delayed: Use the expressions in the following activities

55-78	3 min	Prepare for a reply letter 2: Fill in the blanks. For each blank, half a sentence is provided. The other half (the blank) is to be filled in by translating the Chinese that follows the blank.	T→SS, T→S, SS→T, S→T	• Fill in the blanks with Chinese+ picture prompts • Questioning ("Understand me?") • Gesturing for answers	**Confirm**: "Yes," repeating words from students **Praise**: "Good," "very good" **Awareness-raising**: You should remember "feel like doing something" not "instead of" or "want to."	Delayed: Use the expressions in the following activities
79-106	3 min	Prepare for a reply letter 3: Make up advice sentences. Replaces Chinese prompt with picture prompt.	T→SS T→S S→T SS→T	• Use the focus vocabulary to make suggestions based on picture prompt • Questioning ("Yes?") • Request ("Once again") • Hinting for more answers with rising intonation ("Get up and …?")	**Confirm**: "Yes," repeating words from students **Praise**: "Good," "Perfect," "Very good sentence" **Awareness-raising**: "I think another word is very useful *remind*"	Delayed: Use the expressions in the following activities
107-127	5 min	Prepare for a reply letter 4: Group discussion to provide advice to Li Hua	S→S, T→S, S→SS, T→SS	• Group discussion for advice to Li Hua • Observation	**Praise**: thumb up, "Very good idea" **Awareness-raising**: "You should use as many expressions on the blackboard as possible"; "'If you	Delayed: Use the expressions in the writing activity

				• Group sharing with class	feel' and 'instead of' is frequently used here, right?" **Confirm**: "Yes"; repeating words from students: "Sleep early. Ok" **Further probing**: "Not to sleep. How?" "Go to bed early. That's the same, right?"	
128-140	12 min	Writing the reply letter	T→SS T→S	• Writing on worksheet • Observation	**Reminder**: "Write the answer on the right part. Right, not left" **Focus on form**: "Add a comma here"	Use the expressions taught
141-150	2 min	Assessing the reply letter 1: Introducing assessment rubric + teacher modeling	T→SS SS→T	Questioning: "It's used properly, yes?" "How many points?"	NA	NA
151-158	2 min	Assessing the reply letter 2: Self-assessment and peer-assessment	T→SS S→S	T: Observation SS: Read own letter and other letters in group	**Task management**: "Exchange yours with your partner and then choose whose work is the best"	NA
159-217	7 min	Assessing the reply letter 3: Whole class sharing of "the best" piece of writing	T→S T→SS S→SS	S: read aloud T: questioning: "Why do you think his is the best?"	**Confirm:** "Yes"; repeating words from students **Recast:** "That means most expressions are covered"	Homework

Redirecting: "Good handwriting. Now we should assess it in terms of the vocabulary or words"

Encouraging: "Don't be nervous"

Eliciting student learning

One of the most crucial steps in formative assessment is to elicit student understanding or learning. This is the step in formative assessment that most resembles a traditional assessment tool such as a test or a quiz. However, this elicitation tool is much more varied in shape than our traditional conception of a test. In effect, any way of eliciting student understanding and learning, be it precise or vague, formal or informal, from a look into the students' eyes to a formal achievement test, can be used formatively to inform the teacher and the learners about student learning.

In this lesson, whole-class questioning and individual student questioning were the most often used tool to see if the students understood what was being taught. Table 8 identified 12 types of questioning occurring 84 times in this lesson. Most of the questions asked were management questions such as "Anybody else?" (26). There were also many questions (14) for clarification (e.g., "I feel like sleeping more, yes?"), understanding checks (13) (e.g., "Is that clear?"). Open-ended questions (e.g., "Can you give me some suggestions?") that require answers at some length were also featured highly (12 times).

Fifteen classroom tasks such as group discussion, finding text to read aloud, and individual writing, plus the teacher's observations helped the teacher gauge if the students were learning the four vocabulary items. There were also 16 requests or directions (e.g., "Now, pick it out, please").

Table 8. *Elicitation, interpretation, and feedback practices in this lesson*

Element of CBFA	Type	Technique	Frequency
Elicitation	Questioning (84 times)	Asking about focus vocabulary	1
		Asking for content	2
		Asking for judgment	3
		Asking for translation	1
		Management	26
		Open-ended questioning	12
		Paralinguistic signal	2
		Questioning for clarification	2
		Questioning for confirmation	14
		Questioning for understanding	13

			Rhetorical questions	5
			Rising intonation	3
	Classroom tasks (15 times)		Filling in blanks with prompt	1
			Group discussion	1
			Group rep sharing with class	5
			Making suggestions based on picture prompt	1
			Reading aloud	2
			Reading for peer-assessment	1
			Re-reading for self-assessment	1
			Scanning for info in textbook	1
			Writing on worksheet	1
	Requests or directions (16 times)			16
Interpretation	Information focused (23 times)		Locating information	16
			Providing information	7
	Language focused (39 times)	Accuracy	Targeted vocabulary chunks	28
			Other vocabulary	5
			Other language features	4
		Fluency		0
		Complexity		2
		Appropriateness		0
	Rubric focused (4 times)			4
	Student interpretation (11 times)		Peer-assessment	10
			Self-assessment	1
Feedback	Person-referenced (25 times)		Encouraging	1

		Non-verbal	1
		Praising	23
Task-referenced (29 times)		Awareness raising	6
		Confirming	19
		Focusing on form	1
		Further probing	2
		Recast	1
Management (3 times)		Redirecting	1
		Reminding	1
		Task management	1

Interpretation practices

After data were elicited in this lesson, the teacher made three types of interpretations on the spot. Some focused on information (23 times); most focused on language (39 times); there were also a few times (4) towards the end of the lesson when the teacher focused on a pre-defined assessment rubric for the self-assessment and peer-assessment tasks. In addition, students were involved in interpreting as well during the self-/peer-assessment tasks (Table 8).

Interpretation is the carrier of a teacher's pedagogical content knowledge. In other words, through interpreting learners' understanding and performance, the teacher reveals his or her understanding of what language competence means and how best language is learned. In this lesson, the teacher made sure that the four target vocabulary chunks were repeated in every classroom task, and each task was given enough scaffolding and support. However, when information was the focus, she was mostly asking her students to locate the information in the textbook rather than providing their own information in open-ended language use. In almost all other activities, she interpreted the learners' performance in terms of the accuracy of use for the four vocabulary chunks and a few other words and sentences. In other words, in her interpretation of her students' learning, having learned the four target vocabulary chunks meant the ability to find them in textbook uses and the ability to use them correctly in guided exercises.

Feedback practices

Feedback is the next important step that makes an assessment event formative. In this lesson, the most common feedback provided by the teacher after information was elicited and interpreted included confirming (19 times) and evaluative praising (23 times). Confirming usually took the form of "Yes," "That's right," or repeating what the student had just said. Praising was mostly done verbally, e.g., "good," "perfect," "wonderful," "very good sentence," with an occasional non-verbal thumb-up. In addition, feedback that focused on form was mainly done through raising the students' awareness by commenting on the

importance of the target vocabulary items. Very few instances of corrective feedback were found. Even the only "recast" was not a recast of an error. It was simply saying the same thing in another way.

While confirmations and praises serve to motivate students, they will not achieve the formative function if they do not help close the gap between the learning target and the current level of understanding or performance just elicited. Nevertheless, it should be noted that this lesson was not a new lesson. The teacher confirmed with us later that even the four vocabulary items targeted were not entirely new, and that this lesson was a meaning-focused practice lesson. In this sense, the target of learning was using rather than knowing the four items, and the students were therefore not making many accuracy errors. An occasional minor error was either deliberately ignored or missed in the on-going classroom discourse.

Follow-up actions

The final step that completes one cycle of formative assessment is action. In other words, the feedback stage only informs the teacher and the learners as to what needs to be done. What really makes it formative is actually the follow-up action. In this lesson, follow-up activities were very limited, probably because of the fact that the teacher was targeting practice and use of the vocabulary items, rather than presenting the individual forms as new items. The only instance that can be regarded as a follow-up activity was a planned activity to let students write the letter of advice using the targeted vocabulary.

3.3.4   Discussion

Major components of CBFA are described in the previous section. It is however worth noting that classroom assessment practices do not equate to formative assessment (Black & Wiliam, 2005). We can examine the issue from three perspectives: purpose, practice, and effect. In other words, ideal CBFA should have clear formative purposes, contain all components of at least one formative cycle, and achieve the desired effect. What we analyzed in the findings section were assessment practices. Ideally, for these practices to be identified as formative assessment, they should have been chosen and employed for formative purposes and should achieve formative effects. If we argue that formative effect as an ideal feature does not constitute the minimum defining features of formative assessment, an assessment practice should preferably be aimed at a formative purpose. That said, the transient nature of contingent CBFA makes it hard to determine the momentary purpose which the teacher may or may not be conscious of. In this sense, the teacher's own emic reflection and analysis of these assessment practices are more insightful than the researcher's post-hoc interpretation.

That said, analyzing only classroom assessment practices without teacher involvement can be useful as well. In this lesson, for example, we can clearly see that the teacher was constantly eliciting student understanding, although elicitation was mainly confined to questioning. She was also eliciting student learning through observing her students' performance of classroom tasks. She was constantly providing feedback as well, although most of the feedback was confirming or praising. These types of feedback allowed a smooth running of classroom tasks so that focusing on meaning became possible. It also boosted the students' confidence in using English. On the other hand, the lack of corrective feedback seems to indicate the teacher's vague awareness of the learning target. For instance, why

were these four phrases targeted? For each of the four items, what was being targeted? Form, meaning, or use? Understanding, accuracy, fluency, complexity, or functional use of the four multiword units? Corrective feedback and the entailing formative action should be different for these different targets of learning. A closer look at this lesson suggests that the classroom tasks (elicitation) seemed to be focused on meaning, use, and fluency. However, feedback was mainly targeting accuracy.

In addition, our analysis suggests a lack of complete formative assessment cycles in this lesson. Very often questioning and observations led to quick confirmation or praising. Very few follow-up actions could be found. For complex cognitive growth such as learning, not much can be achieved with even one complete cycle of formative assessment. Furthermore, the usefulness of formative assessment often hinges on a complex web of short-, medium-, and long spiraling cycles (Wiliam, 2010) of monitoring and change. Classroom assessment practices that stop at quick evaluative feedback do not go much further than that in their effectiveness.

4   Summary and conclusion

This article has introduced core features of classroom-based formative assessment, discussed how CBFA research questions should look like, and used a video-taped lesson to illustrate how research questions can be asked, answered and the findings interpreted. We see CBFA as spiraling sets of classroom procedures that are used to elicit students' current level of understanding and performance, to interpret the evidence, provide feedback, and design follow-up activities to close the gap between the current level and the desired level. This entails an understanding that contextualizes CBFA as an integral part of teaching and learning, which in turn suggests that teachers may already be making use of formative assessment in their classrooms. Bringing this practice to a conscious level of awareness will make CBFA more systematic and intentional. A video-taped lesson was used to show how researching CBFA can be done and why this research is beneficial.

Despite the insights we have shown in analyzing the assessment practices in the recorded lesson, this article has revealed an inherent problem of classroom research done by university-based researchers. Assessment practices can be recorded and analyzed, but the intentions and purposes for employing these practices, and the on-the-spot judgment and interpretations of student understanding and learning are by nature subjective and situated. As integral parts of formative assessment, these subjective processes can only be inferred to a limited extent by the external researcher. In this sense, teachers are in the best position to reflect on their own beliefs, behaviors and professional growth. We therefore applaud the model for teacher engagement in the FAR-funded research, and call for more extensive teacher involvement in teacher research and teacher-researcher collaboration. As the primary agents of change in education, teachers' active participation in educational research not only ensures the success of a research project, but also provides the best platform for teacher professional development (Cochran-Smith & Lytle, 1999; Stremmel, 2007).

References

Andrade, H. L. (2010). Summing up and moving forward: Key challenges and future directions for research and development in formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 344–351). New York: Routledge.

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. (2005). Classroom assessment is not (necessarily) formative assessment (and vice‑versa). *Yearbook of the National Society for the Study of Education*, *103*(2), 183–188. https://doi.org/10.1111/j.1744-7984.2004.tb00054.x

Black, P., & Wiliam, D. (2012). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and Learning* (2nd ed., pp. 206–230). London: SAGE Publications.

Cochran-Smith, M., & Lytle, S. L. (1999). The teacher research movement: A decade later. *Educational Researcher*, *28*(7), 15–25. https://doi.org/10.3102/0013189X028007015

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education: Principles, Policy & Practice*, *6*(1), 101–116. https://doi.org/10.1080/09695949993026

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, *43*(3), 393–415. https://doi.org/10.1002/j.1545-7249.2009.tb00242.x

Heritage, M., & Heritage, J. (2013). Teacher questioning: The epicenter of instruction and assessment. *Applied Measurement in Education*, *26*(3), 176–190. https://doi.org/10.1080/08957347.2013.793190

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, *28*(1), 4–13. https://doi.org/10.1002/bs.3830280103

Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, *37*(1), 15–24. https://doi.org/10.1016/j.stueduc.2011.04.003

Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment*, *11*(3), 205–235. https://doi.org/10.1002/tea.20163

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, *44*(1), 57–84. https://doi.org/10.1002/tea.20163

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/BF00117714

Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford, UK: Oxford University Press.

Stremmel, A. J. (2007). The value of teacher research: Nurturing professional and personal growth through inquiry. *Voices of Practitioners*, *2*(3), 1–9.

Torrance, H., & 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. https://doi.org/10.1080/01411920120095780

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18–40). New York: Routledge.

Dr Peter Gu is Associate Professor at the School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand. His research interests include language testing and assessment, learner autonomy and learning strategies, and vocabulary learning. Email: peter.gu@vuw.ac.uk

Dr Guoxing Yu is Professor of Language Assessment and Director of Centre for Educational Assessment and Evaluation at the School of Education, University of Bristol. His research efforts have focused on various aspects of language assessment, and assessment research methods (e.g., eye-tracking). Email: Guoxing.Yu@bristol.ac.uk