

# Journal of Educational Research and Innovation

---

Volume 5

Number 2 *INNOVATIVE ASSESSMENT AND  
TEACHING PRACTICES IN MULTIPLE  
SETTINGS*

Article 4

---

2016

## A Case Study of Formative Assessment to Support Teaching of Reading Comprehension for English Learners

Jane Robin Shore

Mikyung-Kim Wolf

Margaret Heritage

Follow this and additional works at: <http://digscholarship.unco.edu/jeri>

 Part of the [Education Commons](#)

---

### Recommended Citation

Shore, Jane Robin; Wolf, Mikyung-Kim; and Heritage, Margaret (2016) "A Case Study of Formative Assessment to Support Teaching of Reading Comprehension for English Learners," *Journal of Educational Research and Innovation*: Vol. 5 : No. 2 , Article 4.  
Available at: <http://digscholarship.unco.edu/jeri/vol5/iss2/4>

This Article is brought to you for free and open access by Scholarship & Creative Works @ Digital UNC. It has been accepted for inclusion in Journal of Educational Research and Innovation by an authorized editor of Scholarship & Creative Works @ Digital UNC. For more information, please contact [Jane.Monson@unco.edu](mailto:Jane.Monson@unco.edu).

# A Case Study of Formative Assessment to Support Teaching of Reading Comprehension for English Learners

*Jane Robin Shore*  
*Educational Testing Service*

*Mikyung-Kim Wolf*  
*Educational Testing Service*

*Margaret Heritage*  
*WestEd*

Reading comprehension and related processes form the foundation for most of the academic work one does in school. To support the teaching and learning of reading skills for English learners (ELs), teachers seek guidance, resources and intervention programs (Callahan, 2013; Gándara, Maxwell-Jolly, & Driscoll, 2005; Rivera, Moughamian, Lesaux, & Francis, 2008; Teachers of English to Speakers of Other Languages [TESOL], 2010; Walqui & Heritage, 2012). Educators express a need for understanding how to measure EL reading comprehension progress, tailor successful reading instruction, support academic language and content learning, and support overall student success (TESOL, 2010). Reading assessments for formative purposes are argued to be a promising approach for addressing the challenges teachers face, and supporting EL reading in particular (Heritage, 2008, 2012).

Though formative assessment is well supported in the literature, there has been little empirical work done on its use with ELs (Bailey & Carroll, 2015; Black & Wiliam, 1998, Heritage, 2008; Gijbels, & Dochy, 2006; Umer & Omer, 2015; Walqui &

Heritage, 2012; Wei, 2010). This paper describes the design, development, and piloting of a formative measurement system to support EL reading comprehension in the middle grades, the English Learner Formative Assessment (ELFA) system. In particular, this paper reports on a small-scale usability study where the ELFA system was used by eight middle school teachers supporting ELs. We aim to share the lessons we learned during the development and trial of ELFA formative assessment materials for future development and effective implementation of formative assessment for ELs. Our specific research questions and research design are described in the Current Study section.

## Relevant Literature

Formative assessment can be viewed as part of an instructional *process*, where teachers gather evidence of students' learning through assessment during instruction and adapt their instruction to address students' needs. That is, formative assessment is not a test instrument itself (Black & Wiliam, 1998; Heritage, 2010;

Herman, 2013). The formative assessment process is involved with dynamic roles for both teachers and students (Black & Wiliam, 1998). Formative assessment is also described as assessment *for* and *as* learning as opposed to *of* learning (Bennett, 2010).

These qualities of formative assessment may be particularly beneficial for ELs who have diverse needs from their heterogeneous backgrounds. First, identifying learning goals involves, for example, knowledge of individual language proficiency (Solano-Flores & Trumbull, 2008), prior content knowledge (Scarcella, 2002, 2003), and background characteristics (Abedi, 2004), which may present unique opportunities for ELs. Formative assessment uses evidence to drive instruction, which aids in individualized pacing and instructional differentiation (August & Shanahan, 2006; Gándara et al., 2005; Rivera et al., 2008; Walqui & Heritage, 2012). The descriptive nature of feedback on learning used in formative assessment is also particularly effective for ELs, as grades and tests may be culturally-normed and less meaningful for diverse groups (Durán, 2008). Finally, formative assessment captures areas of learning to inform instructional lessons for ELs in real time, like natural language samples, that can be used to identify the strengths and needs of students more accurately (Black & Wiliam, 1998; Heritage, Kim, Vendlinski, & Herman, 2009).

Despite the potential benefits of formative assessment for ELs, little empirical evidence is available to support its development or use (Alvarez, Ananda, Walqui, Sato & Rabinowitz, 2014; Kingston & Nash, 2012; Santos, Darling-Hammond & Cheuk, 2012). One of the reasons might be the challenges in its implementation.

Previous studies have provided empirical evidence with regard to challenges for teachers in the use of effective formative assessment (Bailey & Heritage, 2008; Heritage et al., 2009; Heritage, Walqui, & Linquanti, 2013; Wylie & Heritage, 2010). These include the high demand on teachers' skills and the lack of time to carefully plan and execute effective formative tasks and processes. All teachers have these responsibilities, but there are additional demands placed on those who teach ELs. For example, middle grade teachers typically have scant to no coursework on English language and literacy development or pedagogy (Gándara et al., 2005). Moreover, while setting goals based on learning progression models is fundamental to formative assessment (Heritage, 2008), models specific to ELs are still evolving (Callahan, 2013).

Adding to these challenges, there are few resources for the implementation of effective formative assessment for teachers of ELs. This provides a disincentive for teachers to make formative assessment a part of their practice (Heritage, 2010; 2012). This study was designed to respond to this need, and to contribute to the empirical research on the use of a formative assessment system to support middle school EL reading comprehension.

## Overview of the English Learner Formative Assessment (ELFA) System

Federally funded with a research grant, ELFA was developed and piloted with and for middle school teachers serving ELs as an assessment system for formative purposes. It was also designed to serve as a template or architecture for further individualized development of formative assessment tasks.

### The ELFA Architecture and its Components

ELFA is designed to support intermediate and advanced ELs' reading comprehension of academic argumentative texts and to provide teachers with information to guide instruction. It includes targeted learning goals, assessment learning activities, and teacher support resources. We describe the major components of the ELFA system below.

**ELFA design framework.** A framework document was developed to inform teachers of the specific construct and subskills that were intended to be measured in the ELFA assessment (see Wolf, Shore, & Blood, 2014). The explicit, written description of the construct and subskills was intended to help teachers interpret student responses and understand the gap between the current status and the next step needed for each student. It began

with a description of basic and higher-order reading skills and includes subskills found to be differentially influential in EL reading comprehension (August, Francis, Hsu, & Snow, 2006; Gottardo & Mueller, 2009; Lesaux & Kieffer, 2010; Proctor, Carlo, August, & Snow, 2005; Wong-Fillmore & Snow, 2000). As the focus is specific to the comprehension of argumentative text, the overall approach was also guided by one of the Common Core State Standards in English Language Arts: "Delineate and evaluate the argument and specific claims in a text, including the validity of the reasoning as well as the relevance and sufficiency of the evidence" (National Governors Association for Best Practices and Council of Chief State School Officers, 2010, p. 60). See Figure 1 for the subconstructs and subskills measured in the ELFA assessments.

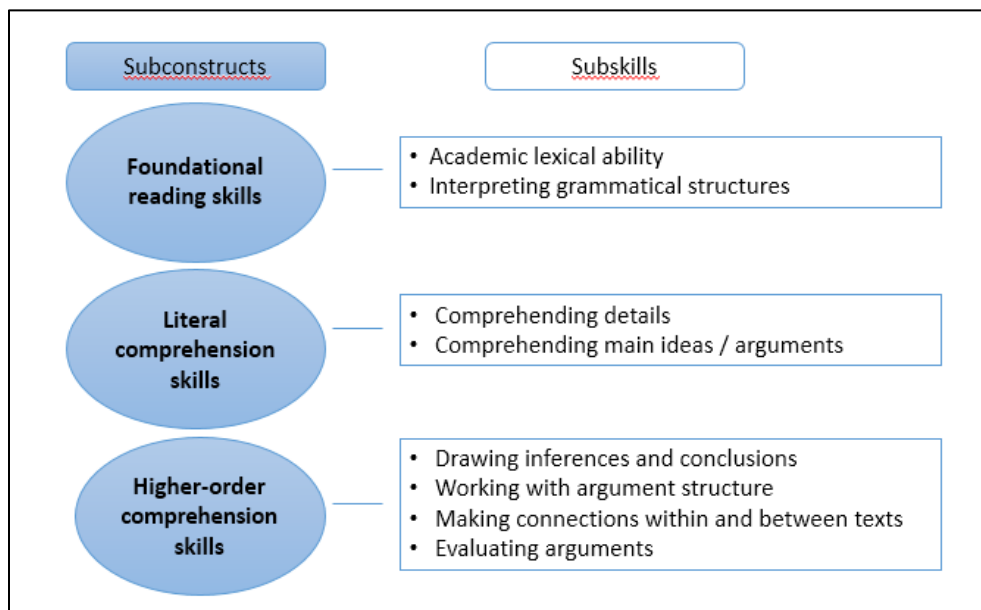


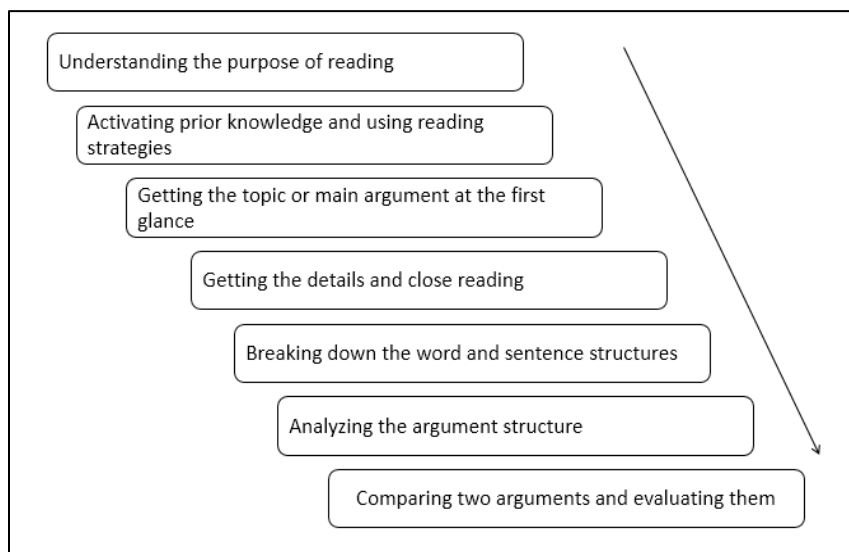
Figure 1. ELFA subconstructs and subskills. Adapted from "Formative Assessment as a Means to Improve Teaching and Learning for English Learners," by M. K. Wolf and J. R. Shore, 2014, Paper presented at the ETS Research Forum. Copyright 2014 by Educational Testing Service.

**ELFA assessment forms.** The ELFA system includes a set of nine reading assessment forms that teachers can use over the course of their instruction. These nine forms are divided into three difficulty categories, *developing*, *intermediate*, and *experienced*, based on the linguistic complexity<sup>1</sup> of the articles in each form. In each assessment form students engage with one main persuasive reading article and a shorter article presenting a counterargument.

Each assessment form also consists of two parts, both based on the same two reading articles and covering the same constructs and subskills. Part 1 contains a

collaborative set of activities and items, and Part 2 includes an individual set. The collaborative tasks that comprise Part 1 were designed to be completed with a peer or in a small group. They were also designed with a purposeful sequence, scaffolded to allow ELs to unpack the given passage and sequentially utilize basic to higher-order reading comprehension skills (see Figure 2).

Teachers interact with students during Part 1 to collect evidence of reading comprehension skills (see Figures 3 and 4 for sample tasks). During Part 2, students work on tasks individually to demonstrate the same skills independently.



*Figure 2.* Task sequencing in ELFA. Adapted from “*Formative Assessment as a Means to Improve Teaching and Learning for English Learners*,” by M. K. Wolf and J. R. Shore, 2014, Paper presented at the ETS Research Forum. Copyright 2014 by Educational Testing Service.

<sup>1</sup> ELFA developers utilized readability software called *e-rater* and *TextEvaluator* to measure dimensions of the linguistic complexity of the passages (Sheehan, 2012; Sheehan, Kostin, & Napolitano, 2012). These tools provided developers with a profile of the linguistic complexity of each reading passage (e.g., the total number of words, lexical density, number of academic words, complexity of sentence structures, grade-level difficulty indices). All reading passages were also rated by focus groups of ESL teachers at the middle-school level for appropriateness of topic, interest, relevance, and language complexity for their students and feedback was provided on which were most relevant, engaging and appropriate for each level. For more information see Wolf, Shore, & Blood, 2014.

**Before you read...**

Look at the article by Sofia Fletcher. Look at the title, the images, and the first sentence of each paragraph.

Where should cell phones be allowed? Where should they not be allowed? What is your opinion? Discuss with your partner and make two lists.

Places where cell phones should be allowed:

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_

Places where cell phones should not be allowed:

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_




Figure 3. Task sample: Warm-up. Adapted from *English Learner Formative Assessment (ELFA), Form 6* by Educational Testing Service. Copyright 2013 by Educational Testing Service.

**Directions:** Read each sentence in the ANSWER CHOICES. One of the sentences is the **main idea** of the article. Three of the sentences are **other ideas** in the article. Two sentences are **not in the article**. Write letters in the blanks in the table below to show where each sentence belongs.

ANSWER CHOICES		
A. Cell phones with internet access can be expensive.		
B. Cell phones should be allowed in schools, but with clear and strict rules.		
C. 63 percent of all students send text messages every day.		
D. Cell phone trucks are a new service that can be found near some schools in New York City.		
E. Students, teachers and parents have different thoughts about cell phones in schools.		
F. Cell phones became very popular in the 1990s because of advertising.		

Main Idea	Other Ideas	Not In The Article
2 _____	3 _____ 4 _____ 5 _____	6 _____ 7 _____

Figure 4. Task sample: Getting a main idea. Adapted from *English Learner Formative Assessment (ELFA), Form 6* by Educational Testing Service. Copyright 2013 by Educational Testing Service.

**Teacher Versions of the assessments.** All ELFA forms include teacher versions (see Figure 5). These are the student forms accompanied by notes and specific guidance intended to support the integration of tasks and teacher interaction

during Part 1 in the form of probing and drill-down questions. General screening questions are provided, as well as guidance on how to drill down to uncover students' thinking and confirm understanding of Part 1 tasks.

**ELFA** English Learner Formative Assessment Getting the Details Student Booklet – p. 5

Directions: Decide whether the sentences in the table are correct or incorrect based on the information in the article.

- If a sentence is correct, circle Correct.
- If a sentence is incorrect, circle Incorrect **AND** quote the article to explain why it is incorrect.
- Two examples have been done for you.

SENTENCES	ANSWERS
<p><b>Example</b></p> <p>An increase in screen time means an increase in time spent being inactive.</p>	<p>Correct    Incorrect</p>
<p><b>Example</b></p> <p>People who spend a lot of time in front of screens consume the same amount of candy and soda as other people.</p>	<p>Correct    Incorrect <i>20</i></p> <p>The article says: "people who spend more time in front of screens consume more high-calorie foods, such as candy and soda."</p>
<p><b>3</b></p> <p>People who spend more time in</p>	<p>Correct    Incorrect</p>

**ELFA** English Learner Formative Assessment Items 3 — 4, p. 10

**Item 3**  
Initial Question Probe:

(1) Why did you choose that answer?

If students cannot find the sentence(s) in the text (paragraph 3) that gives evidence that this is a correct statement or if students do not give a clear explanation that draws on knowledge of key words, go to drill down question(s).

**Drill Down Question:**

(1) Did you use clues in the text and in the question to help you answer this question? What were they?

**Teacher Considerations:**

(1) **Word level:** Do students understand the meaning of the word "sedentary" in the question text? This adjective is key in understanding the effects of screen time described in the sentence. (There is information about the word sedentary in paragraph 3, sentence 2.)

(2) **Sentence level:** Do students point out the particular sentence, "More screen time means less time to prepare healthy foods" (paragraph 3, sentence 6) which gives direct evidence that the statement in the question prompt is correct?

(3) **Discourse level:** Are students able to find the relevant article text to locate the information needed to answer the question? Can students compare the information in the article text with the information in the question text?

Figure 5. Sample item in teacher version. The left part adapted from *English Learner Formative Assessment (ELFA), Form 8* by Educational Testing Service. Copyright 2013 by Educational Testing Service. The right figure adapted from the ELFA Form 8, Teacher Version by CRESST. Copyright 2013 by CRESST/UCLA.

## Current Study

### **Purpose of the Study and Research Questions**

The purpose of this collective case study was to examine the extent to which the ELFA system and its materials were usable and useful for the intended, formative purpose. Specifically, we posited the following research questions:

1. Value: Did ELFA add value to EL instruction?
2. Efficiency: Was ELFA a good use of planning and instructional time?
3. Feasibility: Was ELFA a feasible system for use in classes supporting ELs?
4. Learnability: To what extent were teachers able to learn to use the ELFA system?
5. Professional Knowledge Building: What teacher learning took place while using ELFA?

For the present study, we focused on eight teachers' use of Part 1 of the system. Part 1 of ELFA provided opportunities to observe teacher/student interactions, collaborative peer work, and the use of ELFA to inform and guide instruction; that is, Part 1 focused on all aspects of formative assessment.

### **Methods**

We employed a case study approach in order to closely examine how teachers use ELFA materials as part of their regular instruction in classrooms. Our case study may be described as a *collective* case study, defined as a case study that focuses on a collection of cases to both examine trends that emerge and identify differences between them. As Black and Wiliam (1998) indicate, pedagogy can be very different across contexts that support formative systems. The fact that EL classrooms come with such a range of backgrounds and needs makes a research study necessary.

**Teacher-participants.** A total of eight teachers from six middle schools participated in this study. Among them, seven were ESL teachers and one was an English language arts (ELA) teacher with both ELs and non-ELs in her classroom. All had bachelor's degrees in secondary education and ranged in teaching experience, having taught both ELs and non-ELs between five to 30 years. All were teaching in urban school districts and supported ELs with intermediate to advanced English language proficiency (ELP) at the time of data collection. Some teachers supported mixed level classes

(e.g., intermediate and advanced in the same classroom) while other teachers had a homogeneous class of students with either intermediate or advanced ELP.

Participating 6<sup>th</sup>, 7<sup>th</sup>, and 8<sup>th</sup> grade classes were chosen based on the similarity of their ESL programs. Students' ELP levels were determined by each state's annual summative ELP assessment. Most students were long-term ELs while some students (8%) were newcomers. Table 1 describes each of the teachers' settings, experience, learners served, and languages to further contextualize the case study settings.

Table 1

*ELFA Participating Teachers, Students and Programs*

Teacher ID	Subj	State	School EL size	# of years teaching	# of ELs	Grades	Student ELP levels	Home languages spoken by most ELs	Other languages spoken by teacher
T1	ESL	NJ	4%	23	35	7,8	Adv	Spanish Haitian	Spanish
T2	ESL	NJ	2%	12	12	6,7,8	Int	Creole	None
T3	ESL	NJ	12%	11	41	6,7,8	Int-Adv	Spanish	Spanish
T4	ESL	NJ	3%	12	5	7	Int-Adv	Polish	Spanish
T5	ELA	CA	48%	15	48	8	Adv	Spanish	None
T6	ESL	CA	48%	8	13	7,8	Int	Armenian	None
T7	ESL	OR	74%	30	19	6	Int-Adv	Spanish	None
T8	ESL	OR	74%	5	9	6	Int	Spanish	None

Note that because the unit of study was the teacher, we focused our analysis on the teacher and the overall classes - not the individual students.

**Study instruments.** Interviews were conducted during training and before and after each ELFA use session, based on a set of interview protocols. The questions were guided by the research questions. An

observation protocol was also used to systematically document the details of all sessions in which ELFA was used in classrooms. The protocol included taking detailed notes of the teacher's introduction of the ELFA forms, classroom discourse, teacher interaction with students, and the use of the teacher's version of ELFA, including the probing questions. For further



examples of specific items, see Figures 3 and 4 above.

**Procedure.** The participating teachers were first provided with all of the ELFA materials for review prior to training. Next, teachers attended three small group (two or three teachers) webinars intended to introduce the ELFA system and provide guidance on how to use it for formative assessment purposes

A week after the training, each teacher used the ELFA assessment forms and teacher versions over the course of two weeks (two to four lessons for each teacher). At least two researchers observed each lesson. Teacher interviews took place before and after each lesson, and were recorded and transcribed.

**Analysis of data.** A coding scheme was developed to analyze the observation notes and interview transcripts. Following the procedure suggested by Miles and Huberman (1994), a pair of researchers performed a preliminary round of coding while making detailed notes. The initial coding and memos were discussed among the researchers and the coding scheme was refined. The pair of researchers conducted multiple readings of the observation notes and interview transcripts to reach consensus on their codings based on the final coding scheme. Table 2 summarizes the coding scheme that was applied to the study's data.

Table 2

*Coding Scheme for the Usability of ELFA*

Dimensions	Description & Subcategories
Useful	Comments on...
1) Value	• ELFA as a tool for collecting learning evidence for ELs
2) Efficiency	• Interpretability of learning evidence
	• Adaptability for lesson planning
	• Any areas of improvement in the ELFA materials
Usable	
3) Learnability	• Teachers' use of various ELFA materials (assessment forms, teacher versions including probing questions and observation guidance)
4) Feasibility	• Use of integration during regular instruction
	• Alignment with standards and curricula
	• Any practical constraints in using ELFA
Other Themes (Professional Knowledge)	• Teacher understanding of formative assessment practice
	• Teacher perception on reading comprehension skills for ELs

**Results**

Using the coding described in Table 2, several themes emerged from the data to answer our research questions. In addition

to findings about the usability of the ELFA system, other themes emerged, demonstrating the changes in each teachers' thinking.

**Four Dimensions in the Usability of ELFA**

**Value.** Value was defined as whether the teachers perceived the ELFA system as providing them with something they did not already have. Five of the teachers (T1, T3, T4, T6 and T8) commented that ELFA was useful because it was more aligned with their instructional planning than other tests available to them. For two of the teachers

(T1 and T3), discussion of what they liked about ELFA started with an expression of what they did not like about other available materials. For example, as one teacher (T3) noted, current materials available for middle school ELs place too much focus on foundational skills rather than higher-order thinking skills. The results are summarized in Table 3 below.

Table 3

*ELFA Usability: Value*

Expression of Value	Teachers	Comments
Aligned with instructional planning	T1, T3, T4, T6, T8	
Dissatisfaction with other available materials	T1, T3	<p>“I’ve been so dissatisfied with the materials that exist because they challenge [students] to understand grammar and vocabulary but they’re not really doing the academic tasks and enriching activities that are going to reinforce their academic growth in the general sense” (T3).</p> <p>“Part of my problem is that materials appropriate for their skill levels, the vocabulary and the things that are being discussed, are everyday things. The kids have been here in the U.S. for a long time, so everyday tasks are covered ... to get them stepped up to be able to argue and discuss and do academic tasks and use academic language to complete tasks, I think that is something that is different from what we are currently using” (T1).</p>
Better aligned with instructional practice and philosophy	T4	<p>“...these [ELFA] materials are more aligned with what I want my kids [ELs] to be doing, and it’s better than what I’m currently using... They have to demonstrate a lot of ... skills....to have them interact in a way that’s a bit more enriching and in-depth, it’s valuable...” (T4).</p>
Well-designed forms	T6, T8	<p>“This feels really targeted at looking at breaking down subskills and trying to look critically at it” (T6)</p> <p>“If I were to have this kind of assessment and the assessment would pinpoint to me the areas of need, I can group the students according to those areas of need at a time” (T8).</p>

In summary, five of the participating teachers specifically remarked on the value in the ELFA system as an approach that expanded their current resources and their understanding of skills, in addition to being a classroom-based assessment system that could readily be integrated into instructional planning.

**Efficiency.** Efficiency was defined as whether teachers found the use of ELFA to be an efficient use of their time. Many of the participating teachers' concept of efficiency evolved over the case study period. At the start, four of the teachers worried aloud about the time it took to use ELFA in the classroom (T1, T4, T5 and T8). During the pre-observation interviews, one of these teachers (T4) said, "I think there's incredible amount of opportunity in ELFA for collaboration and cooperative learning ... I was really interested in something, but we didn't have time and had to move on to the next one!" In the initial lessons, participating teachers tended to focus on "getting everything done." They were watchful of the time and anxious about completing items and the activities. Views changed over time.

During observations, and through interviews, it was found that teachers began to emerge with a more of a focus on the *process* of learning, and less on the "right" answer. During observations, researchers noted that five of the eight teachers emerged as more and more oriented toward the activity than the timing (T1, T2, T4, T5, and T8). This was seen in actions like encouraging learners to talk about the article in their own words, to confirm their understanding with peers, and develop a way to express why they might choose certain answers in activities. This finding is further supported explicitly in interviews; for example, after the second

lesson of ELFA integration, one teacher (T4) expressed, "I was freaking out about the time, but what I should have done is just get to work ... to see if they can do it."

After the second period using ELFA one teacher (T5) said,

"Once I read it, I just was concerned about time. We move *really* slow at these levels...but as we moved through, the ELFA highlights areas and we can pick and choose. I shouldn't be worried about speed, but about their real understanding..."

Overall, initial concern about the efficiency of integrating ELFA into their instruction was resolved through the flexibility in the use the assessment. It was also helpful for teachers to discuss during interviews how to use ELFA after each lesson to better understand the intent.

**Learnability.** Learnability was defined as how easily and smoothly teachers were able to work the materials into their teaching practice. Upon reviewing the codings of this dimension, teachers were found to fall into three groups: "knowledge building," "emerging awareness," and "evolving practice" groups. Results are summarized in Table 4 below.

The coding results from the learnability dimension suggest that teachers were generally able to integrate ELFA into their teaching practice, but with varying levels of ease.

**Feasibility.** As far as the implementation of a new program or strategy such as ELFA is concerned, its feasibility in the given context is an important consideration. In this collective case study, feasibility was defined as the degree to which ELFA would be a good fit in the current curriculum teachers were responsible for delivering to their students.

As seen in the results for efficiency and learnability, some variation was noted in the teachers' perceptions of feasibility. The results for feasibility are summarized in Table 5 below

Teachers' perceptions of feasibility appeared to be partly dependent upon individual teachers' pedagogical practices and interpretations of how ELFA would be used with their existing curricula.

Table 4

*ELFA Usability: Learnability*

Group	Teachers	Comments
Knowledge Building	T1, T2, T8	"I wondered if it's important that [students] do this right now or have them get on with the task...that was me struggling and figuring out what I should be doing" (T1).
Emerging Awareness	T3, T4, T5	"It would be helpful for there to be an option to have somebody come out and say "this is how it looks," and everybody's together and we're doing it all together as a group" (T3).  "I should have spent more time with the ELFA materials before this went down...sat have with the team and learn the system better" (T4).
Evolving Practice	T6, T7	I think that's something that...is helpful and easy to do. I'm actually pretty excited about that" (T6). "Next week I'm continuing with ELFA, so if these materials became available to me on a website then I would be quite capable of getting the materials and incorporating them into instruction" (T7).

Table 5

*ELFA Usability: Feasibility*

Group	Teachers	Comments
Feasible	T5, T6	". . . working together, implementing strategies, and having the time, we could definitely do this. ...So I don't think it's too difficult" (T5).  "I think it would be relatively simple. I can do this given the support materials, and colleagues" (T6).
Evolving	T4	"Some students found it hard to work collaboratively...they wanted to go at their own pace ... There was real guidance, though, so...it is a good learning. It fits our goals" (T4).

### **Other Emerging Themes**

Besides usefulness and usability, additional themes emerged across with our eight teacher participants. These spanned changes in teachers' conceptualization and articulation about formative assessment to evolution of their understanding of ELs' reading comprehension skills. The results we present in this section include these themes in assessment thinking that emerged upon coding of the observation notes and interview transcript data.

**Changes in the understanding of formative assessment.** During the debriefing interviews after each of the observation days, teachers had opportunities to discuss their perspectives on the use of ELFA and ask clarifying questions. The cycle of pre-interview, lesson, and post-interview seemed to serve as professional development on the use of ELFA for the teachers. One notable theme was the participating teachers' thinking around the intended purpose of ELFA as a formative assessment system. While all teachers were in support of formative assessment conceptually, four teachers (T2, T4, T5 and T7) explicitly expressed that they could understand the basic tenets of formative assessment.

However, during the first observations, researchers found that all of the teachers' talk during circulation among the students focused mainly on clarifying the directions of ELFA tasks. The majority of the teachers

also encouraged students to not dwell on answers and move on in order to complete the ELFA assessment. In all but one of the classes the researchers observed on the first day, teachers neither used the teacher version of the assessment nor asked any probing questions to confirm understanding as to why students chose certain answers.

That said, by the third or fourth day of integration, observation notes began to reveal changes in teacher thinking among six of the eight teachers. Six teachers clearly began to express a change in their concept of formative assessment (T1, T2, T5, T6, T7, and T8). Some teachers also talked about looking at the suggested probing questions for each item from the teacher version the night before the observation (T3 and T4). They commented that the teacher version helped them get their thoughts in order for upcoming lessons. Although observations did not reveal instances of using probing questions in all cases (four of the eight teachers, T1, T2, T5 and T8, used them during observations), interviews revealed changes in teacher thinking among all teachers, which we might predict would influence future lessons. Further, five teachers (T1, T2, T3, T5, and T7) also began to discuss the focus on higher-order thinking in various ways, demonstrating an understanding of formative assessment as a progression of learning. These findings are summarized in Table 6 below.

Table 6

*Other Emerging Themes*

Theme	Teachers	Comments
Change in concept of formative assessment	T1, T2, T5, T6, T7, T8	<p>“I need to slow down, ask questions and look for clues in what students do and say to help pinpoint challenges. I feel like it’s making me think differently about assessment. It’s way more powerful than a score” (T7).</p> <p>“What I would do differently next time is do it with them and then repeat the activity over and over again with different documents as we move forward ... And I would keep working in different ways until they would be able to get the process” (T8).</p>
Focus on higher-order thinking	T1, T2, T3, T5, T7	“The information we get here, it’s about learning and informing my job here. It’s just terrific, really. It’s so surprising now, refreshing. It’s not about a score, but it’s still an assessment” (T5).
Use of deep probing and confirmation of understanding	T1, T3, T4, T7	<p>“Today was such an eye opener when [students] had to identify subjects and verbs in complex sentences! In my mind I was thinking, “How could you not know this?” I wouldn’t have realized I needed to go back there” (T1).</p> <p>“This really supports my decision making and plans for the week. For example, tomorrow I will work on those word families. They clearly had trouble with those...” (T3).</p>
Focus on wide range of subskills	T1, T3, T5, T7, T8	“It’s useful data that I can scan through. I can pick and choose where to go based on evidence in items from all the students...” (T7).

The observations and teacher discussions indicate that some teachers were accustomed to thinking about assessment only in the context of summative assessments and scores. As an aspect of formative assessment driving instruction, the use of evidence is a clear indicator of changes in teacher thinking about assessment. The researchers looked for indications that teachers might be

delving deeply into student thinking and confirming understanding for the purpose of informing future instruction. During the initial lessons with ELFA, the researchers observed that all teachers had the tendency to move through lessons quickly, spending little time on items if students got the correct answers. However, observations painted a very different picture after several days of ELFA use and reflection.

Examples of using evidence to inform thinking or even modify instruction were clear in all but one of the classrooms. In half of the classes (T1, T3, T4 and T7), the teachers began probing deeply and confirming understanding. They were speaking of using this information to make instructional planning based on evidence. As indicated in these comments, teachers began to take time to think about possible next activities based on what they observed during their use of the ELFA system.

**Changes in the understanding of reading comprehension skills and instruction.** Another notable theme lay in teachers' expanded views of the underlying sources of reading comprehension difficulties for ELs that may be attributed to the use of ELFA. At the start of their collaboration with the ELFA team, most teachers described the major source of ELs' reading comprehension difficulty in terms of deficiency in vocabulary knowledge. Seven of the eight participating teachers mentioned that a major focus of their instruction supporting ELs was on words. Most described similar sentiments to T4, who said:

"We are always working on vocabulary. They get the words, we go online, find definitions, and memorize definitions. And there is *nothing* else you can do. When you learn a second language, you memorize...that's the only way you can learn it."

As the teachers were using the ELFA assessment forms, their teaching and activities in class began to change. For example, one teacher (T3) described,

"The paraphrasing is great. I did have to remind them that paraphrasing was putting that same idea in your own words. A lot of them were a little off. But in doing this, I realized that they

were getting the main ideas and vocabulary to do so. It wasn't only about building vocabulary. It was about building the ideas and concepts that formed text."

Observations also provided some insight into teachers' conceptual understanding of reading comprehension and how ELFA might have been seen to influence it. While using ELFA, all teachers were asking questions, probing about the difference between the topic, main idea, and main argument. One of the teachers (T4) held a class discussion on the similarities between the main idea and the main argument. During the observations, three teachers (T2, T4 and T5) dove deeply into students' conceptual knowledge about text. During the post-observation interviews, one teacher (T2) reflected by saying, "What I would do is have them do a paragraph by paragraph summary. Maybe that would help them locate the main idea vs. main argument. Maybe breaking it into pieces..." Another teacher (T5) said, "I usually teach a lot of pre-vocabulary, but this (ELFA) makes me wonder if instead, the time could be better spent focusing on working on main ideas and concepts, then activities that have students demonstrate their comprehension."

Reflections from teachers and observations support a change in orientation about reading comprehension. Observation notes and interviews reveal that seven of the participating teachers, as described here, went from a focus on words to a focus on a wider range of subskills. Rich reflections provided further insight into what the teachers had learned while using ELFA, and also paths for further ELFA development.

## Discussion

While views about the value, efficiency, learnability, and feasibility of ELFA were generally positive among the participating teachers, it is also worth noting the important areas that teachers mentioned in the context of improving classroom-based assessments and systems for formative assessment purposes. First, it became clear that any formative assessment system needs to be aligned directly with the key areas teachers feel they need to focus on in instruction, or aligned well with concepts or skills specific to their curriculum. Five of the eight participating teachers specifically pointed to the current lack of appropriate materials to be used for formative purposes for middle school ELs' reading comprehension skills, and they highlighted the provision of such material as one of ELFA's areas of strength (T1, T3, T4, T6 and T8). Whereas new academic standards such as the Common Core State Standards have brought higher reading-skill demands for students, it appeared to the teacher-participants in the current study that it is equally important to have both foundational and higher-order reading skill activities and tools to be used for EL students—a perception that is echoed in the current national conversation (see Bailey & Carroll, 2015; Santos et al., 2012).

A second key consideration is that, for any formative assessment system or resource to be an efficient use of time, it must be modular and flexible (T1, T4, T5, T8). By “modular” we mean that formative assessment forms, activities, or tools should contain activities that can be used both individually (to focus on certain skills) or as a group (to cover a variety of skills). Formative systems must also include guidelines for flexible implementation. This flexibility needs to be built not only

into the formative assessment system, but also into teacher development.

Another important finding of the present case study was that even the most experienced teachers' thinking and engagement with formative assessment appeared to change over time while using the ELFA materials. The use of ELFA, and subsequent discussions between the researchers and the teacher-participants, served as professional development for the teachers. Use of formative support systems over time can, we contend, inform practice in a professionally engaging way.

Our finding about the initial mismatch between the teachers' concepts about formative assessment and the actual use of assessments for formative purposes may be partly attributed to the careful planning required for formative assessment. Formative assessment places large cognitive demands on teachers to collect, interpret, and act upon evidence quickly (Sondergeld, Bell, & Leusner, 2010). As Heritage et al. (2009) indicate, once teachers begin collecting such evidence, they also must learn new ways of pacing, differentiating, organizing, and adapting their instruction. This learning takes time and is not always internalized automatically.

### **Limitations**

As a case study, the findings of this research are not intended to generalize to a larger population of teachers or students, but rather to serve as an illustration of how integration of a new system like ELFA might work across a few settings. The purpose is to provide a deeper understanding of the cases presented, and not a larger perspective of the use of assessments for formative purposes. That said, there were a few limitations in the study.

First, our cases were limited to those in urban environments. This was purposeful,



as we aimed to ensure that our teacher-participants had relevant and extended experiences with ELs. We described these settings with details intended to aid teachers reading this article to decide how relevant the work might be for their own settings.

This study was not an experimental study to examine the effectiveness of ELFA on teaching and learning. Rather, it was an exploratory study, conducted over a short period of time, to investigate the potential usability of ELFA for formative purposes in classrooms with EL students. That said, the short time period means that the study provides a snapshot of the use of ELFA rather than a picture of use over an extended period of time, such as a school year.

The study also focused exclusively on the teachers' practices and perspectives on formative assessment at this particular ELFA development stage—not on those of their learners. The observed behaviors and expressed perspectives of our educator partners were the unit of study, although student interactions may have influenced these areas.

## Conclusions and Future Research

Despite the limitations, the study yielded useful insights and support for future development work and an expanded research agenda for ELFA. Perhaps the most important finding was that teachers not only found value in the system, but they also felt that its use improved their knowledge and practice. Questions worth further exploration are: What longitudinal changes emerge in teachers who support ELs when they integrate formative processes? What other factors facilitate the effective integration of ELFA into the classroom to carry out systematic formative

assessment practice? A longitudinal study can offer valuable suggestions to effectively implement formative assessment using a system or tools such as ELFA. Further, a future study should include an investigation of student perspectives on learning and EL students' reading outcomes in classes where the ELFA is used regularly. The impact on EL students' reading comprehension outcomes as a result of ELFA use would be a useful and important contribution to the field.

**Jane Robin Shore, Ed.D.** is a Research Scientist in the Center for English Language Learning and Assessment at the Educational Testing Service. She can be contacted at [ishore@ets.org](mailto:ishore@ets.org).

**Mikyung Kim Wolf, Ph.D.** is a Senior Research Scientist at the Center for English Language Learning and Assessment at the Educational Testing Service. She can be contacted at [mkwolf@ets.org](mailto:mkwolf@ets.org).

**Margaret Heritage, Ph.D.** is a Senior Scientist with Program Services at WestEd. She can be contacted at [mherita@wested.org](mailto:mherita@wested.org).

## Acknowledgements

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A100724 to ETS. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The authors would like to thank Terry Cryan, Ian Blood, Kristin Williamson, Jennifer Wain, Barbara Jones, Sandy Chang and Jennifer Lentini for their contributions to this research project.

## References

- Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, 33, 4–14.
- Alvarez, L., Ananda, S., Walqui, A., Sato, E., & Rabinowitz, S. (2014). *Focusing formative assessment on the needs of English language learners*. San Francisco, CA: WestEd.
- August, D., Francis, D., Hsu, H-Y. A., & Snow, C. (2006). Assessing reading comprehension in bilinguals. *The Elementary School Journal*, 107(2), 221–238.
- August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language Minority Children and Youth*. Mahwah, N.J: Lawrence Erlbaum.
- Bailey, A.L., & Carroll, P.E. (2015). Assessment of English learners in the new era of content standards. *Review of Education Research*, 39(1), 253–294.
- Bailey, A. L., & Heritage, M. (2008). *Formative assessment for literacy, grades K-6: Building reading and academic language skills across the curriculum*. Thousand Oaks, CA: Corwin Press.
- Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning: A preliminary theory of action for summative and formative assessment. *Measurement: Interdisciplinary Research and Perspectives*, 8, 70–91.
- Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139–148.
- Callahan, R. M. (2013). *The English learner dropout dilemma: Multiple risks and multiple resources* (California Dropout Research Project Report #19). Santa Barbara, CA: University of California.
- Durán, R. P. (2008). Assessing English-language learners' achievement. *Review of Research in Education*, 32, 292–327.
- Educational Testing Service (2013). *English Learner Formative Assessment (ELFA), Form 8*. Princeton, NJ: Author.
- Gándara, P., Maxwell-Jolly, J., & Driscoll, A. (2005). *Listening to teachers of English language learners: A survey of California teachers' challenges, experiences, and professional development needs*. Sacramento, CA: The Regents of the University of California.
- Gijbels, D., & Dochy, F. (2006). Students' assessment preferences and approaches to learning: Can formative assessment make a difference? *Educational Studies*, 32(4), 399–409.
- Gottardo, A., & Mueller, J. (2009). Are first and second language factors related in predicting L2 reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, 101, 330–344.
- Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment*. Washington, D.C.: Council of Chief State School Officers.
- Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Washington, D.C.: Council of Chief State School Officers.

- Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice*, 28(3), 24-31.
- Heritage, M., & Chang, S. (2012). Teacher use of formative assessment data for English language learners. Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Heritage, M., Walqui, A., & Linqunti, R. (2013, May). *Formative assessment as contingent teaching and learning: Perspectives on assessment as and for language learning in content areas*. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.
- Herman, J. (2013). *Formative assessment for next generation science standards: A proposed model*. Princeton, NJ: K-12 Center at ETS. Retrieved from <http://www.k12center.org/rsc/pdf/herman.pdf>.
- Kingston, N., & Nash, B. (2012). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28–37.
- Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, 47, 596–632.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage Publications.
- National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA (2013). *English Learner Formative Assessment (ELFA), Teacher's Version*. Los Angeles, CA: Author.
- National Governors Association for Best Practices and Council of Chief State School Officers (2010). *Common Core State Standards: English language arts and literacy*. Washington, D.C.: Authors.
- Proctor, C. P., Carlo, M. S., August, D., & Snow, C. E. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, 97, 246–256.
- Rivera, M. O., Moughamian, A. C., Lesaux, N. K., & Francis, D. J. (2008). *Language and reading interventions for English language learners and English language learners with disabilities*. Portsmouth, NH: Research Corporation, Center on Instruction.
- Santos, M., Darling-Hammond, L., & Cheuk, T. (2012). *Teacher development to support English language learners in the context of Common Core State Standards*. Palo Alto, CA: Stanford University, Understanding Language Initiative.
- Scarcella, R. (2002). Some key factors affecting English learners' development of advanced literacy. In M. J. Schleppegrell and M. C. Colombi (Eds.), *Developing advanced literacy in first and second languages* (pp. 209–226). Mahwah, NJ: Erlbaum.
- Scarcella, R. (2003). *Academic English: A conceptual framework* (Tech. Rep. No. 2003-1). Irvine, CA: University of California, Linguistic Minority Research Institute.

- Sheehan, K.M. (2012). *A cognitively-based text analysis system designed to help test developers ensure that admissions assessments incorporate suitably complex text*. Riyadh, Saudi Arabia: National Center for Assessment in Higher Education.
- Sheehan, K. M., Kostin, I., & Napolitano, N. (2012, April). *SourceRater: An automated approach for generating text complexity classifications aligned with the Common Core Standards*. Paper presented at the National Council on Measurement in Education, Vancouver, BC.
- Solano-Flores, G., & Trumbull, E. (2008). In what language should English language learners be tested? In R. J. Kopriva (Ed.), *Improving testing for English language learners: A comprehensive approach to designing, building, implementing, and interpreting better academic assessments* (pp. 169-200). New York: Routledge.
- Sondergeld, T. A., Bell, C. A., & Leusner, D. M. (2010). Understanding how teachers engage in formative assessment. *Teaching & Learning: The Journal of Natural Inquiry and Reflective Practice*, 24, 72–86.
- Teachers of English to Speakers of Other Languages (TESOL). (2010). *TESOL/NCATE standards for the recognition of initial TESOL programs in p-12 ESL teacher education*. Alexandria, VA: TESOL.
- Umer, M. & Omer, A.M.A. (2015). An investigation of Saudi English-major learners' perceptions of formative assessment tasks and their learning. *English Language Teaching*, 8(2), 109-115.
- Walqui, A., & Heritage, M. (2012, January). *Instruction for diverse groups of ELLs*. Paper presented at the Understanding Language Conference, Stanford, CA.
- Wei, L. (2010). Formative assessment: Opportunities and challenges. *Journal of Language Teaching Research*, 1(6), 838-841.
- Wolf, M. K., & Shore, J. (2014, March). *Formative assessment as a means to improve teaching and learning for English learners*. Paper presented at the ETS Research Forum, Washington, D.C.
- Wolf, M. K., Shore, J. R., & Blood, I. (2014). *English Learner Formative Assessment (ELFA): A design framework*. Princeton, NJ: ETS.
- Wong Fillmore, L. & Snow, C. (2000). *What teachers need to know about language*. Washington, D.C.: ERIC Clearinghouse on Languages and Linguistics. Retrieved from [http://www.elachieve.org/images/stories/eladocs/articles/Wong\\_Fillmore.pdf](http://www.elachieve.org/images/stories/eladocs/articles/Wong_Fillmore.pdf).
- Wylie, E. C., & Heritage, M. (2010). Developing and deepening formative assessment practice. In M. Heritage (Ed.), *Formative assessment: Making it happen in the classroom* (pp. 117–132). Thousand Oaks, CA: Corwin Press.