ELPA Validity Evaluation Instrument:

Interview protocol for content-area teachers regarding the English language proficiency of students in their classroom

Washington
Idaho
Indiana
Montana
Oregon

edCount, LLC
Center for Assessment
UCLA
Synergy Enterprises, Inc
PIRE

March 2011

About EVEA

The U.S. Department of Education funded Enhanced Assessment Grant *Evaluating the Validity of English Language Proficiency Assessments* (EVEA; CFDA 84.368) was awarded to the Office of the Superintendent for Public Instruction of the State of Washington in fall 2009. The project brought together five states – Idaho, Indiana, Montana, Oregon, and Washington – to work on collaborative and independent validity plans for English language proficiency assessments (ELPAs) over an 18-month period. During the EVEA funding period, none of the partner states belonged to an existing ELPA consortium; rather each had worked with commercial test developers to create state-wide ELPAs that are aligned with their state English language development (ELD) standards. The main project goal was for each state to create a validity argument for its ELPA system. Additional project outcomes included:

- Building individual State Interpretive Arguments for the validity of each state's ELPA,
- Building a Common Interpretive Argument for any ELPA;
- Designing a set of studies and instruments to support and pilot test these arguments; and
- Making instruments publically available at the close of the project for the wider education community to access.

This research instrument is one product of these efforts.

Collaborating institutions

edCount, LLC
The National Center for the Improvement of Educational Assessment (NCIEA, the Center for Assessment)
The Graduate School of Education and Information Studies (GSE&IS) at the University of California, Los Angeles (UCLA)
Synergy Enterprises, Inc. (SEI)
The Pacific Institute for Research and Evaluation (PIRE)

Authors

This validity evaluation instrument was prepared by:

Mari Quenemoen, ID Research Partner

Allison Kerbel, IN Research Partner

Research Objective

Performance levels should meaningfully reflect the range of English language proficiency that students placed within each level demonstrate. Weakness in the appropriateness of performance levels can have direct consequences for students, since most placement and exit decisions are based on the performance levels into which students test. This interview protocol for content-area teachers can be used to investigate the appropriateness of both the performance levels into which students are placed and the cut score used to determine (in whole or in part) a student's readiness to exit from EL services. Content-area teachers' observations of students' strengths and areas of development in writing, reading, listening, and speaking may reveal the accuracy and appropriateness of the performance levels into which students test in the four domains and overall. A weak alignment between teachers' observations of students' skills and students' actual ELPA performance levels could indicate the need to review either the ELPA standard-setting process or the quality of the performance level descriptors.

This document outlines the claims, underlying assumptions, and research questions that can serve as the foundation for potential/future studies related to this topic. In addition, it provides sample questions for an interview protocol to be used with content-area teachers who have ELs in their classrooms.

Claim

ELPA performance levels reflect meaningful differences in students' English language proficiency.

Underlying Assumptions

Cut scores are set so that performance levels meaningfully reflect of the range of English language proficiency that students placed within each level demonstrate.

Performance level descriptors (PLDs) accurately describe the range of linguistic skills that students who test into the performance level demonstrate.

The performance level(s) associated with the exit decision reflect a level of proficiency at which language skills are no longer a barrier to participation in academic discourse in the classroom.

Research Questions

Research Question 1: To what extent do the performance levels into which students are placed meaningfully reflect these students' range of English language proficiency?

Research Question 2: To what extent do students who have exited the ELPA system possess the speaking, listening, reading, and writing skills and sufficient content-specific academic vocabulary to participate and make progress in the content area classroom?

Method

The research questions listed above can be addressed by conducting interviews with content-area teachers of ELs about their observations of students' strengths and areas of development in academic English language in the classroom. If teachers' observations of students' relative language development do *not* align with the students' most recent performance levels on the ELPA, this lack of alignment may indicate a weakness in the appropriateness of the cut scores. This section outlines the rationale for this approach and includes an exemplar interview protocol.

The chosen research question will determine the sampling method. To investigate the appropriateness of performance levels (Research Question 1), it is important to find teachers who teach a number of EL students at widely divergent levels of English language proficiency. To investigate the top cut score that informs exit decisions (Research Question 2), it is important to find teachers who teach a number of ELs who have been placed in the top performance level and a number of ELs who recently exited from EL services. In preparation for the interviews, a list of students in each participating school whose ELPA scores placed them into each performance level will be generated, along with a list of teachers who have at least two of these students in their class. In order not to influence teachers' assessments of students' abilities, teachers should not be told how each student scored on the ELPA.

The study should include at least 25 students and at least 8 teachers. If the initial study indicates that there may be a problem with the cut scores, the study could be expanded. Choosing teachers from multiple grade levels may reveal whether scores are better aligned with teachers' observations at some grade levels than at others. It is also important to collect information about teachers' level of experience, both as classroom teachers and specifically teaching ELs, to gauge whether this affects the quality of information that teachers provide about EL students' language skills.

Analysis

Research Question 1:

Teachers' observations of students' skills in each domain (listening, speaking, reading, and writing) can be compared with students' domain-level performance levels (if the state sets performance levels for each domain) or scores that students received for each domain on the ELPA. Teachers' predictions of students' overall ELPA performance levels can be compared with students' actual performance levels. After asking teachers to read the performance level descriptors carefully, a poor alignment between teachers' predictions of students' ELPA performance levels could indicate one of a number of weaknesses: weakness in the cut scores that divide the performance levels, weakness in the accuracy of the performance level descriptors themselves, or weakness in the design of the test. The alignment of teachers' observations in the domains and students' domain performance levels could indicate weaknesses in the cut scores for the individual domain performance levels, or a weakness in the design of questions measuring that domain. Teachers' observations of students' control of content-area vocabulary may provide supplemental information to explain students' performance in the classroom, though this may not be tested on the ELPA.

Research Question 2:

Teachers' observations of students who scored "fluent" or "proficient" on the ELPA should demonstrate a readiness to participate in content-area discourse in the classroom. If teachers observe that exited or proficient students still struggle with the academic English needed to participate fully in the classroom,

the top cut score may have been set too low. Conversely, if students who scored below the top cut score demonstrate a level of English proficiency that poses no barrier to full participation in the classroom, the top cut score may have been set too high.

Interview Protocol

Demographic Information

- For how many years have you been a teacher?
- Which grades have you taught?
- What is your educational background (i.e., degrees, content areas, special certification)?
- Have you ever received any specific training or professional development in English language development or instruction for English learners? If so, what?

Interview Questions

Questions	Analysis	Construct Evaluated
1. Is ___(name)___ able to comprehend academic content that is conveyed verbally during class? A)Always; B) Most of the Time; C) Some of the time; D) Rarely Can you give me a specific example?	Compare with student's listening performance level. Identify and analyze common themes from additional information.	Listening
2. Do you think ___(name)___ is able to express verbally [his/her] understanding of academic content well enough to participate in academic discussions? A)Always; B) Most of the Time; C) Some of the time; D) Rarely Can you give me a specific example?	Compare with student's speaking performance level. Identify and analyze common themes from additional information.	Speaking
3. Do you think ___(name)___ understands written English well enough to comprehend selected readings in your class? A)Always; B) Most of the Time; C) Some of the time; D) Rarely Can you give me a specific example?	Compare with student's reading performance level. Identify and analyze common themes from additional information.	Reading

4. Do you feel ___(name)___ is able to <u>express</u> [his/her] understanding of academic content in <u>writing</u>? A)Always; B) Most of the Time; C) Some of the time; D) Rarely Can you give me a specific example?	Compare with student's writing performance level. Identify and analyze common themes from additional information.	Writing
5. Compare ___(name)___ with the native English speaking students in your class. How are this student's challenges and areas of development in <u>writing</u> similar or different from [his/her] native English speaking peers?	Identify and analyze common themes from additional information.	Writing
7. How often is the student's <u>basic content area vocabulary</u> sufficient for the student to express his/her content knowledge? A)Always; B) Most of the Time; C) Some of the time; D) Rarely Can you give me a specific example?	Identify and analyze common themes from additional information. Code for specific content area.	Content-area vocabulary
8. Read carefully these ELPA performance level descriptors. Into which performance level do you think ___(name)___ scored on the most recent ELPA? OR How do you predict that ___(name)___ will score on the next ELPA?	Compare with actual student performance levels; calculate percentage of matches; calculate whether teachers over- or under-estimated more frequently for each performance level.	Overall