

# Performance-Based Assessment: Rubrics, Web 2.0 Tools and Language Competencies

*W. I. Griffith, Defense Language Institute Foreign Language Center<sup>1 2</sup>*

*Hye-Yeon Lim, Defense Language Institute Foreign Language Center<sup>3</sup>*

## Abstract

The problem of assessing student performance in foreign language classrooms is one of the biggest tasks facing educators. Assessments based on actual performance, clear measurement rubrics and the use of technology have all been shown to increase student motivation and performance. This paper examines ways in which performance-based assessments (PBAs) can help teachers assess actual student competencies in all language skills. The paper begins with a general discussion of how to create specific types of rubrics and presents sample rubrics designed to measure student competencies in various skills. Tasks are developed to measure two specific competencies required by SEP and ways in which free Web 2.0 tools can be used with these rubrics are discussed.

## Resumen

El problema de valorar desempeño de los estudiantes en aulas de idioma extranjero es una de las tareas más importantes a las que se enfrentan los docentes. Se ha demostrado que las evaluaciones basadas en el desempeño real, las rúbricas claras y el uso de la tecnología aumentan la motivación de los estudiantes así como su desempeño académico. Este artículo examina diferentes maneras en las que evaluaciones basadas en el desempeño (PBAs) pueden ayudar a los maestros a valorar las competencias reales de los estudiantes en todas las habilidades del idioma que están aprendiendo. El escrito empieza con una discusión general de cómo crear tipos específicos de rúbricas y presenta muestras de rúbricas diseñadas para medir las competencias del estudiante en varias habilidades. Se desarrollan tareas para medir dos competencias específicas que la SEP requiere y se discuten las maneras en que se pueden usar las herramientas gratuitas de Web 2.0 con estas rúbricas.

## Introduction

Assessing student performance in ESL/EFL classrooms is one of the biggest concerns educators face. What does it mean to give students a grade? Does a passing grade mean they really have communicative proficiency in the language? The increasing emphasis on competency means that educators must devise methods that measure proficiency and the ability to perform.

---

1 Contents of the paper are not necessarily the official views of or endorsed by the U.S. Government, Department of Defense, Department of the Army, or the Defense Language Institute Foreign Language Center.

2 [wanda.griffith@dliflc.edu](mailto:wanda.griffith@dliflc.edu)

3 [hyeyeon.lim@us.army.mil](mailto:hyecheon.lim@us.army.mil)

Educators must find ways to measure whether a student can use information to confront real-world tasks successfully.

Many teachers have traditionally relied on some sort of test to assess learning. The problem with this approach is that while tests may assess how much information a student has retained, they do not often evaluate how well a student can use this knowledge to perform a task. Multiple-choice exams, for instance, make it difficult to measure language competency demanding more than recall of the subject matter (Brualdi, 1998, Roediger, 2005). The emphasis on the assessment of competencies demands a new way of thinking about how to evaluate students and requires varied forms of assessments to determine the extent to which students can actually use knowledge to complete tasks.

The purpose of this paper is to discuss how to evaluate student competency and proficiency in foreign language classrooms. We begin with an examination of performance-based assessments (PBAs). Next, we discuss how to build a rubric for specific skills based on a performance task. Finally, we show how some of the new Web 2.0 tools can be used to carry out performance-based assessments using specific rubrics. We will use competencies from SEP (Secretaría de Educación Pública, México – Ministry of Public Education) to illustrate the ideas discussed.

### **Performance-Based Assessments (PBAs)**

It has often been noted that while students may perform well on a test, they have no actual command of the language and cannot use it in authentic situations. Measuring the ability to use the language in real-world situations requires putting the student into a situation where s/he produces language to complete an authentic task (Lim & Griffith, 2011). Performance-Based Assessments have been very useful for this approach.

PBAs “represent a set of strategies for the...application of knowledge, skills, and work habits through the performance of tasks that are meaningful and engaging to students” (Hibbard et al., 1996, p.5). Such assessments provide teachers with information about how well a student understands and applies knowledge. It goes beyond the ability to recall information and beyond rote memorization of rules.

Performance activities are often based on “authentic” tasks. In order to be “authentic”, the National Capital Language Resource Center (NCLRC, 2011) notes that assessment activities must meet the following criteria:

- Be built around topics or issues of interest to the students;
- Replicate real-world communication contexts and situations;
- Involve multi-stage tasks and real problems that require creative use of language rather than simple repetition;
- Require learners to produce a quality product or performance;
- Evaluation criteria and standards are known to the students;
- Involve interaction between assessor (instructor, peers, self) and person assessed; and

- Allow for self-evaluation and self-correction as students proceed.

Many benefits of PBAs have been identified (cf. Blaz, 2001; Tedick, 1998). The following is a list of these benefits:

- Performance tasks engage and interest students;
- PBAs are accurate and meaningful indicators not only about what students know but about what they can do;
- PBAs can increase student confidence because students know and understand the standards;
- PBAs can improve clarity because, by sharing the rubric, students know exactly what is expected to get a certain grade or score;
- PBAs increase teacher confidence in assessing student learning; and
- Performance tasks let teachers know how well they are teaching and let students know how well they are learning.

In sum, performance-based assessments are based on real world tasks and measure whether students can apply the knowledge and skills to accomplish these tasks. The benefits are that students are more engaged and have more information about what they need to do than is the case with more traditional assessments. Finally, because performance tasks are generally taken over multiple points in time, they can provide a more accurate assessment of what a student knows and what a student is able to do.

#### *Types of Performance-Based Assessments*

Performance-based assessment, as the name implies, measures how well a student actually performs while using learned knowledge. It may even require the integration of language and content area skills (cf. Brualdi, 1998; Valdez Pierce, 2002). The key is the determination of how well students apply knowledge and skills in real life situations (Frisby, 2001; McTighe & Ferrara, 1998; Wiggins, 1998). The successful use of PBAs depends on using tasks that let students demonstrate what they can actually *do* with language.

There are three types of performance-based assessment from which to choose: *products*, *performances*, or *process-oriented assessments* (McTighe & Ferrara, 1998). A *product* refers to something produced by students providing concrete examples of the application of knowledge. Examples can include brochures, reports, web pages and audio or video clips. These are generally done outside of the classroom and based on specific assignments.

*Performances* allow students to show how they can apply knowledge and skills under the direct observation of the teacher. These are generally done in the classroom since they involve teacher observation at the time of performance. Much of the work may be prepared outside the classroom but the students “perform” in a situation where the teacher or others may observe the fruits of their preparation. Performances may also be based on in-class preparation. They include oral reports, skits and role-plays, demonstrations, and debates (McTighe & Ferrara, 1998).

*Process-oriented assessments* provide insight into student thinking, reasoning, and motivation. They can provide diagnostic information on how

well students use learning strategies and may lead to independent learning when students are asked to reflect on their learning and set goals to improve it. Examples are think-alouds, self/peer assessment checklists or surveys, learning logs, and individual or pair conferences (McTighe & Ferrara, 1998).

### **The Rubric**

Actually determining a student's competency at a given task, however, requires more than a good activity. It requires a device or instrument to measure the skill. Typically, this tool is a rubric. A rubric includes the specification of the skill being examined and what constitutes various levels of performance success. Constructing an appropriate rubric is central to meaningful performance-based assessment.

#### *Steps in Constructing a Rubric*

Constructing a good rubric involves focusing on exactly what to measure, how to measure performance, and deciding what a passing level of performance competency is. Although a general rubric design can be used multiple times, it may have to be fine-tuned or altered slightly to meet the specific requirements of any given task activity. This section examines the process in more detail.

#### 1. Defining the Behavior to Be Assessed

The first step is to decide what behavior students must perform. The SEP<sup>4</sup> information provides one set of expectations regarding what students need to be able to accomplish at the end of each unit and at the end of each term. Several important questions need to be addressed:

- What concept, skill, or knowledge am I trying to assess?
- What should my students know?
- At what level should my students be performing?
- What type of knowledge is being assessed: reasoning, memory, or process? (Stiggins, 1994)

By answering these questions, you can decide what type of activity best suits your assessment needs.

#### 2. Choosing the Activity

After defining the purpose of the assessment, decide on an activity. Decisions should consider issues regarding time constraints, resources, and how much data is necessary (Airasian, 1991; Popham, 1995; Stiggins, 1994).

#### 3. Defining the Criteria

After determining the activity and the tasks to be included in the activity, define which elements of the project/task will be used to determine the success of the student's performance. Airasian (1991, p. 244) suggests several steps to complete that process:

---

<sup>4</sup> The most recent competencies for English education can be found on the SEP website at <http://www.sep.qob.mx>

1. Identify the overall performance or task to be assessed, and perform it yourself or imagine yourself performing it;
2. List the important aspects of the performance or product;
3. Try to limit the number of performance criteria, so they can all be observed during a student's performance;
4. If possible, have groups of teachers think through the important behaviors included in a task;
5. Express the performance criteria in terms of observable student behaviors or product characteristics;
6. Don't use ambiguous words that cloud the meaning of the performance criteria; and
7. Arrange the performance criteria in the order in which they are likely to be observed.

### *Characteristics of Good Rubrics*

Most rubrics consist of three basic items: objectives, performance characteristics, and points or scores indicating the degree to which the objectives were met. Unlike more traditional tests, performance-based assessments do not have clear-cut right or wrong answers. Instead, there are degrees to which the student is successful or unsuccessful. Rubrics need to take those varying degrees into consideration.

A rubric should be thought of as a rating system to determine the proficiency level at which a student is able to perform a task or display knowledge of a concept. With rubrics, it is possible to define the different levels of proficiency for each criterion. When using any type of rubric, you need to be certain that the rubrics are fair and simple. Also, the performance at each level must be clearly defined and accurately reflect its corresponding criterion or subcategory (Airasian, 1991; Popham, 1995; Stiggins, 1994).

Rubrics need to clearly define what constitutes excellent, good or poor work. Students need to be able to see exact expectations. Similarly, teachers need to be able to recognize those same characteristics. The guidelines should be consistent as well as clear since this gives students some assurance that the teacher will not engage in "subjective" grading (Blaz, 2001).

According to Blaz (2001), good rubrics are identified by several features. First, they contain only observable behaviors and are phrased in positive, rather than negative terms. The markers are specific (e.g., makes *many* errors should be makes *more than ten* errors). Good rubrics do not ask students to do things they have not already been shown or taught. For example, do not ask students to use tenses which have not been taught or to use vocabulary and expressions that are beyond the scope of the unit they are studying or have studied. Good rubrics should allow for a clear demarcation between acceptable/passing behavior and unacceptable/failing performance.

### *Types of Rubrics*

Two types of rubrics are often used in scoring performance: *holistic* and *analytic* (Mertler, 2001; Moskal, 2000). A *holistic* rubric evaluates the overall performance in a qualitative manner. Speaking scores on the iBT TOEFL, for example, are graded holistically. Scores on such scales give an overall impression of student ability and often use a 4 or 5 point scale. One common example is given by the categories "excellent", "good", "fair", and "poor."

*Analytic* rubrics break down the performance into different levels of behavior and assign point values to each. Points are then totaled to derive a quantitative measure of performance. For example, a speaking rubric might include the dimensions of pronunciation, use of proper tense, use of transitions, vocabulary, and fluency. The advantage of such rubrics is twofold. First, it is possible to give different weights to the different dimensions reflecting their relative importance. The second is that they provide more information to students about relative strengths and weaknesses.

It is important to remember that a good rubric must assess not only a concept or skill but also what the students *should* know based on the criteria. Teachers need to be able to assess the level at which students are performing as well as what kind of knowledge (e.g., memory, application, evaluation) is being measured (Stiggins, 1994). It is crucial to ensure that more than simple memory is being assessed because in the real world it is rarely sufficient to have memorized something.

### **Carrying out the Assessment**

There are myriad ways to conduct assessments in foreign language classrooms. The most important thing is to remember that not all activities can be used as performance-based assessments (Wiggins, 1993). Performance-based assessments require the application of knowledge and skills *in context* and not just the simple completion of a task. That is, they generally require higher-order thinking skills (Lim & Griffith, 2011).

Unlike traditional assessment methods where all students complete a test or task at the same time, the idea of PBA is based on the notion that students will have multiple assessments taken over time to determine how the student is performing. This implies that not all students need to be evaluated at the same time although, by the end of the term, all students will have the same number of assessments.

For example, suppose that a teacher wishes to assess oral behavior in a large classroom. Suppose further that the teacher has assigned some oral group work, perhaps an information gap activity. The teacher can wander around the class or even sit in the center of the room and listen to the students as they participate. When a student speaks, teachers can write down a code next to their name. The code could be based on a simple *holistic* rubric. At the end of the week, the teacher looks at the codes and comes up with a weekly score for that student.

Teachers can also assign tasks that need to be turned in to be graded. In this way, all students will submit their work at the same time for evaluation. Similarly, teachers could assign oral reports or group performances which could be evaluated in class.

One common way of assessing student work is through the portfolio. Here, the students keep copies of their work (ideally their best work) and submit them for review at the end of the term. In this way, it is easy to see the progress students have been able to make. Typically, one thinks of such portfolios as being most appropriate for writing (essays, posters, letters, etc.) but with Web 2.0 tools, nearly every kind of skill can have a portfolio.

The assessment of the receptive skills (i.e., listening and reading) is problematic. It is very difficult to do product or performance assessments for these skills without requiring some sort of product output. How does one determine whether or not a student has understood a reading passage unless the student is able to answer some questions about the reading? Still, it is possible to meet this challenge and to generate an e-portfolio using Web 2.0 tools.

With free Web 2.0 tools, one can easily create histories of performance or portfolios for listening, speaking, reading, and writing. Using these portfolios or historical trails, teachers can have a clear record with which to assess a student's ability to apply language in authentic settings. Furthermore, these tools give all students the opportunity to participate more than they could in a classroom. For example, in a classroom of fifty students, individual students can participate in very few oral skills projects during a semester. With the Web 2.0 tools, time is not such an issue since many of the products can be done outside the class and reviewed by the teacher later.

#### Web 2.0 Tools and Assessment Strategies

A variety of tools are now available on the Internet to help teachers do authentic assessment. These free tools are easy to use and can be adapted for in-class or out-of-class tasks. Though it is beyond the scope of this paper to offer a full discussion of these tools, we will mention several with potential for use in ESL classrooms.

VOKI<sup>5</sup>, EYEJOT<sup>6</sup>, and VOXOPOP<sup>7</sup> can be easily used to assess speaking. Each of these has the capacity for students to record their voices and send them to a teacher. They offer the opportunity for students to respond to authentic materials and tasks.

Excellent tools, such as PENZU<sup>8</sup> and TITANPAD<sup>9</sup>, can be used to create writing portfolios. TITANPAD differs from PENZU in the fact that it allows for collaborative work in real time by multiple persons writing on the same pad and identifiable by unique color fonts.

---

<sup>5</sup> <http://www.voki.com>

<sup>6</sup> <http://www.eyejot.com>

<sup>7</sup> <http://www.voxopop.com>

<sup>8</sup> <http://penzu.com>

<sup>9</sup> <http://titanpad.com>

Another great tool for writing is the text-to-speech animated moviemaker XTRANORMAL<sup>10</sup>. Here students can choose from a variety of sets and characters, customize characters, and animate their actions. Students type a dialogue that the program then changes to speech when the movie is created.

WALLWISHER<sup>11</sup> is a tool that allows the creation of virtual “sticky note” boards on which students can place images, audio files, or texts and is a tool where multiple skills can be used. It has the further advantage of allowing students to add to the board over time creating a kind of timeline for teachers to view progress.

### *Two Examples Using Competencies Specified in the SEP Curriculum*

The following examples illustrate how the steps discussed above can be implemented to create a rubric and identify a Web 2.0 tool for assessment for a SEP specified competency.

#### Example 1:

In Level 1, a SEP (Aguirre, 2006) competency requirement specifies that “*Students can use language creatively and appropriately by choosing lexis, phrases and grammatical resources in order to produce short, relevant texts (form letter/e-mail, conversation) regarding factual information of a personal kind.*” How can teachers use some of the Web 2.0 tools to assess this?

This competency can be demonstrated in a variety of ways. Since a key is “short text/factual information,” a self-introduction could be used. This self-introduction can be assessed by using either writing or speaking activities. Students could create a “VOKI” or avatar (graphic image) themselves and then record their introduction and send it to the teacher. Alternatively, students could write an introduction of themselves which VOKI would then change to speech.

Another option might be for students to use XTRANORMAL to create, in pairs, a movie where two people meet for the first time. They would exchange greetings and give personal information. This is likely to be an engaging task and may diminish some of the stress of performing in front of the class. The outcomes provide products that can be kept in a portfolio for comparison with latter ones.

So far, teachers have decided how to *get* a product but *not* how to *evaluate* the skills. As noted earlier, the teacher now has to decide exactly what is being evaluated. There are a number of things that could be evaluated here including pronunciation, intonation, grammar, proper use of lexis, and so on.

Suppose the task requires the student to provide a spoken self-introduction. The student can use VOKI to send a video introducing him/herself to the teacher. The following is a possible *analytic* rubric that might be used to

---

<sup>10</sup> <http://www.xtranormal.com>

<sup>11</sup> <http://www.wallwisher.com>

grade this performance focusing on the amount of personal information, breadth of vocabulary, proper use of language, and fluency. These four categories are equally weighted in this rubric design.

| Factual Information  | Vocabulary   | Correctness of Language  | Fluency  |
|--|--|--|--|
| 3<br>It contained more than 5 facts about the student.     | 3<br>Student used a wide variety of vocabulary words to describe self (more than in the lesson). | 3<br>Less than 10% of words contain pronunciation errors; two or fewer grammar mistakes; 0-1 incorrect word choices.                 | 3<br>Introduction was told with expression, not flat; good L2 intonation pattern; confident in speech.                   |
| 2<br>The student gave between 3-5 facts about him/herself. | 2<br>Student used the necessary vocabulary and basic vocabulary was used correctly.              | 2<br>Between 11% and 30% of words have pronunciation errors; between 3 and 6 grammar errors; between 2 and 4 incorrect word choices. | 2<br>Some expression in speech though mechanical in places; a few pauses but they didn't detract from comprehensibility. |
| 1<br>The student gave less than 3 facts about him/herself. | 1<br>The introduction used words incorrectly and used  | 1<br>More than 30% of words have pronunciation   | 1<br>Very flat presentation; little L2 intonation  |

Table 1: An *Analytic* Rubric Example for a Speaking Task

To assess the ability, a teacher would listen to the video, circle the number in the appropriate boxes in the rubric and sum the scores. The scores range from 4-12. Teachers could then create a grading scale as follows: A= 11-12; B=9-10; C= 7-8; D=5-6, and F=4 or less. These grades could be registered for the students' records and the video presentations kept for portfolio purposes and comparison with later progress.

Example 2:

In Level Three, a SEP requirement states students *"can recognize and understand quotidian texts (diaries, personal notes, letters/e-mails, timetables, diagrams of public transport, road maps, travel brochures/guides, advertisements, plane/train/bus tickets and conversations) in order to use them purposefully (schedule meetings/appointments, get/give prices, locate*

*places, find/propose alternative routes, discuss future plans*) (Aguirre, 2006, p. 99)."

To address this requirement, the teacher could post a map of a city on WALLWISHER or could hand out a map for students to use. The map should include some buildings known to the students so that they can use the buildings for reference and description. Students are asked to specify a place where they would like to meet a friend. The students then write a message with specific instructions using PENZU and mail it to the teacher. The students can, if they want, upload images of the buildings or sites to which they refer in their directions. The following is a sample of a *holistic* rubric that can be used to assess the writing an email containing directions.

| Exceeds Expectations  | Meets Expectations  | Fails to Meet Expectations  |
|---|---|---|
| <ul style="list-style-type: none"> <li>• All directions are in the imperative form.</li> <li>• 90% of the directions were correct.</li> <li>• More than 6 directions are used</li> <li>• Grammatical errors are minor and don't impede comprehensibility.</li> <li>• Includes more than 4 or more physical references with descriptions.</li> </ul> | <ul style="list-style-type: none"> <li>• Most (80%) of the directions are in the imperative form.</li> <li>• At least 75% of the directions were correct.</li> <li>• Includes at least 4 but fewer than 6 directions.</li> <li>• Some grammatical errors but they don't impede comprehensibility.</li> <li>• Includes 2-3 physical references with descriptions.</li> </ul> | <ul style="list-style-type: none"> <li>• Uses fewer than 4 directions.</li> <li>• More than half of the directions are incorrect in form.</li> <li>• Less than 75% of the directions were correct.</li> <li>• Grammatical errors impede comprehensibility.</li> <li>• Used fewer than 2 physical references with descriptions.</li> </ul> |

Table 2: A *Holistic* Rubric Example for a Speaking

In this example, the teacher chooses WALLWISHER to deliver information to students and PENZU to collect the data. The skills being assessed include grammar, use of the imperative form for giving directions, the accuracy of the directions given, the extent of detail provided, and the complexity of the directions as measured by the number of steps (directions). The teacher uses a *holistic* rubric in Table 2 to decide which category best fits the students' emails and to assign one of three grades: "exceeds expectations," "meets expectations," or "fails to meet expectations." Note that this type of rubric gives a general sense of performance that is less detailed than the *analytic* rubric. Remember, this rubric should be given to students. It can be posted on WALLWISHER with the original assignment and instructions so that each student knows exactly what must be done to achieve a certain rating (e.g., include at least four physical references with descriptions). If necessary at some later point, these could be turned into letter grades such as A, C, or F.

## Conclusions

Measuring the skills and knowledge of students in the ESL/EFL classroom can be a cumbersome process. The tactics of using multiple-choice tests or fill-in-the-blank quizzes have not proven to be accurate indicators of how students might perform in real situations. Therefore, it is time to focus more on performance-based assessments that allow the measurement of a student's ability to use the language in authentic situations.

The strength of such assessments depends on the rubrics actually used in the evaluation of the skill. Both *holistic* and *analytic* rubrics have a place in helping to understand skill levels. Good rubrics will focus on a specific skill or task and therefore allow for accurate and meaningful measures of language use.

Finally, the activities designed are also important. New technologies such as Web 2.0 give teachers access to a variety of useful tools that make it easier to create authentic tasks, to require the use of higher order thinking skills, and to go beyond memorizing. These tools also allow the student to use skills in a variety of ways and situations so that language is transferred across situations.

As competency and the ability to use a language effectively become increasingly important, we must develop new ways to improve students' performances and to make learning more engaging, productive, and effective. Creating assessments based on performance, devising clear rubrics, and introducing technology into the classroom have all been shown to improve motivation and performance. By understanding the relationship among performance assessments, rubrics, and activities, we can build classrooms where students can succeed and where what they learn applies to their needs in the real world, in the classroom, and beyond.

## References

- Aguirre, E. M. (2006). *Lengua Extranjera: Inglés, Programas de Estudio 2006*, Mexico City: Secretaría de Educación Pública.
- Airasian, P.W. (1991). *Classroom Assessment*. New York: McGraw-Hill.
- Blaz, D. (2001). *A Collection of Performance Tasks and Rubrics: Foreign Languages*. Larchmont, NY: Eye on Education.
- Brualdi, A. (1998). Implementing performance assessment in the classroom. *Practical Assessment, Research & Evaluation*, 6 (2). Retrieved June 10, 2010 from <http://PAREonline.net/getvn.asp?v=6&n=2>
- Frisby, C. L. (2001). Academic achievement. In L.A. Suzuki, J. G. Ponterotto, & P. J. Meller (Eds.), *Handbook of Multicultural Assessment* (pp. 541-568). San Francisco: Jossey-Bass.
- Hibbard, K. M., Van Wagenen, L., Lewbel, S., Waterbury-Wyatt, S., Shaw, S. & Pelletier, K. (1996). *A Teacher's Guide to Performance-based Learning and Assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Lim, H-Y, & Griffith, W.I. (2011). Practice doesn't make perfect. *MEXTESOL Journal*, Vol. 34 No .3
- McTighe, J., & Ferrara, S. (1998). *Assessing Learning in the Classroom*. Washington, DC: National Education Association.

- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). Retrieved July 4, 2011 from <http://www.jcu.edu/academic/planassess/pdf/Assessment%20Resources/Rubrics/Other%20Rubric%20Development%20Resources/Designing%20Scoring%20Rubrics.pdf>
- Moskal, B.M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, 7(3). Retrieved July 4, 2011 from <http://www.peopledev.co.za/library/Scoring%20rubrics%20-%20Moskal%20B.pdf>
- National Capital Language Resource Center (NCLRC). (n.d.). *The Essentials of Language Teaching*. Retrieved September 5, 2011 from <http://nclrc.org/essentials/assessing/alternative.htm>
- Popham, W. J. (1995). *Classroom Assessment: What Teachers Need to Know*. Needham Heights, MA: Allyn and Bacon.
- Roediger, H.L. III (2005). The positive and negative consequences of multiple-choice testing [Electronic version]. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 31(5), 1155-1159.
- Stiggins, R. J. (1994). *Student-centered Classroom Assessment*. New York: Macmillan Publishing Company.
- Tedick, D. J. (Ed.) (1998). *Proficiency-oriented Language Instruction and Assessment: A Curriculum Handbook for Teachers*. Minneapolis, MN: University of Minnesota, Center for Advanced Research on Language Acquisition.
- Valdez Pierce, L. (2002). Performance-based assessment: Promoting achievement for English language learners. *News Bulletin ERIC/CLL*, Vol. 26 (1). Retrieved from [http://www.cal.org/resources/archive/news/2002fall/CLLNewsBulletin\\_Fa02c.pdf](http://www.cal.org/resources/archive/news/2002fall/CLLNewsBulletin_Fa02c.pdf)
- Wiggins, G. (1993). Assessment, authenticity, context, and validity. *Phi Delta Kappan*, November, 200-214.
- Wiggins, G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco, CA: Jossey-Bass.