

Leadership for Achievement

Balanced Assessment Systems for Linguistically Diverse Schools

Rebecca Freeman Field (rdfield@casloninc.com); Diep Nguyen (nguyend@d62.org)

Improving the achievement of all students, particularly English language learners and students with disabilities, demands an internal assessment and accountability system that yields real evidence of student growth relative to state standards, program goals, and local benchmarks. This session presents the BASIC model, a state-of-the-art system that draws on common formative and summative assessments included in a pivotal portfolio. Participants examine how to use data to inform decisionmaking at the classroom, school, program, and district levels. We highlight practical strategies for using data to foster collaboration among mainstream, ESL, bilingual, and special education teachers, and for planning meaningful professional development.

The Big Idea for School Leaders

Schools must rely on multiple sources of evidence to monitor and organize student learning within local contexts. Therefore, the most efficient and reliable approach is to build a **system of assessments** to help collect both formative and summative data that are valid, reliable, and relevant over time.

What do we know? We know that...

- 1. English language learners are among the fastest-growing demographic group of students in U.S. schools.** Between 1995 and 2006, the ELL student growth rate exceeded 57% while the total pre-K – 12 student growth rate was less than 4%. Most ELLs live in six states (CA, TX, FL, NY, IL, AZ), and 27 other states/territories have over 20,000 ELLs. The density and growth rate of ELLs varies dramatically across states, with over 100% growth in nearly 25 states (including AL, AR, DE, GA, ID, IA, KS, KY, MI, NE, NC, SC, TN, UT, WV).
- 2. There is tremendous diversity in the category “English language learner”.** ELLs vary in terms of first language/cultural background, proficiency in English, socioeconomic status, and educational achievement, and this diversity influences their participation and achievement at school. About 80% of these students are Spanish speakers (and there is great diversity within this category), and the remaining 20% represent over 400 different language groups. The majority of these students are U.S.-born citizens (76% of elementary ELLs and 56% of secondary ELLs).
- 3. NCLB accountability requirements have focused increased attention on the achievement gaps among subgroups in all schools, particularly ELLs.** US schools are held accountable for the academic achievement in English of all students, and for the English language proficiency development of ELLs, as measured by standardized tests.
- 4. NCLB’s approach to accountability is highly criticized** for being overly rigid, punitive, unscientific, and likely to do more harm than good for the students who are now being left behind. Nowhere is this more true than in the case of English language learners (*James Crawford, 2006*).
 - A standardized test that was designed for English speakers to measure their content-area knowledge in English is not a valid or reliable assessment for students who are defined as English language learners.
 - It takes 5 – 7 years or more for ELLs to develop the academic English they need for school success.
- 5. A valid and reliable accountability system includes multiple criteria of student performance and development** (*Crawford, 2006; Gottlieb, 2003*). A broad consensus has emerged among testing experts that achievement tests of questionable validity and reliability – or, indeed, a single test of any kind – should not be used for high-stakes decision-making. Instead, we should be building intelligent, reciprocal accountability systems. (*Darling-Hammond, 2010*).
- 6. Race to the Top (RTTP) highlights the need for data systems** (one of the four pillars of educational reform) **that provide educators with information on student progress to drive their decision-making and ensure that every child can and does achieve in U.S. schools.**

Foundations (from Gottlieb & Nguyen, 2007)

Purposes of Assessment for linguistically diverse learners at the Program/District Level

- Provide multiple sources of evidence of student **growth, proficiency and achievement** in language development (native language /L1 and second language/L2), academic learning (in L1 and/or L2), and socio-cultural competence
- Monitor student group progress to guide instructional, curricular, and program decisions
- Document effectiveness of instructional practices and program implementation for public reporting purposes
- Identify patterns of instructional challenges and create a plan for School/Program/District Improvement Plan

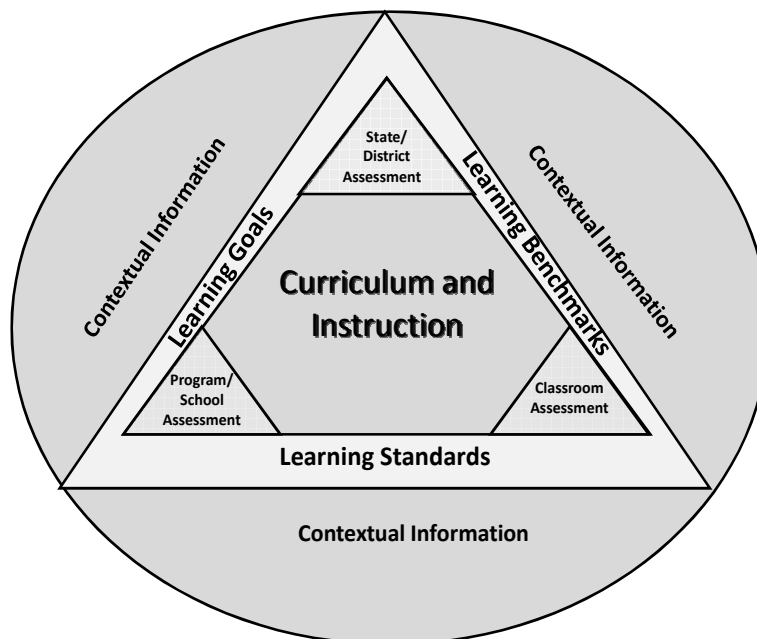
Features of Authentic Assessment and Accountability

- Internal to the functioning of schools and school districts while responsive to external accountability mandates
- Built on consensus from both teachers and administrators
- Rigorous, comprehensive, and standards-based
- Systemic, reflective of shared educational goals, vision, and commitment
- Directly related to teaching and learning

How can a balanced assessment system help establish accountability? By...

- **Building success:** Provide multiple sources of evidence of student growth in language proficiency, academic proficiency and achievement, and sociocultural competence.
- **Documenting success:** Measure and provide evidence of student achievement and instructional effectiveness.
- **Ensuring success:** Help diagnose challenging areas and create a plan for improvement

The BASIC Model
Balanced Assessment and Accountability System, Inclusive and Comprehensive



Gottlieb & Nguyen (2007)

Complement Standardized Tests at the State/ District Levels with Common Formative Assessments at the School/Program Levels:

State/District level

Standardized tests of academic achievement and English language proficiency: *Summative data*

- Used to determine the number of students attaining proficiency in one year
- Responds to federal and state accountability requirements

School/Program level

Standard or common assessments: *Summative and formative data*

- Used to assess student growth and achievement OVER TIME in relation to district/state standards using district rubrics
- Allows for meaningful comparisons across classes, programs, schools in district
- Used to inform teaching, learning, program and professional development, and policy

Classroom level

Idiosyncratic or classroom-based assessments: *Formative data*

- Used to guide instruction
- Used to monitor students' growth and evaluate student achievement relative to teaching and learning objectives

Pivotal Portfolio

A hybrid of the **working portfolio** (students' work-in-progress) and **the showcase portfolio** (students' best work) with three main distinctions:

1. Each teacher gathers what the teachers collectively consider evidence of essential student learning and achievement.
2. All teachers use common assessments of that essential student work.
3. Pivotal portfolio follows the student for the length of the students' career in the language education program.

1 x 1 x 1 One child - One portfolio - One adult

Developing A Balanced Assessment System in Your District/Program:

A Step-by-Step Process

Identifying gaps in your system...

1. Identify federal, state, and local mandates and accountability requirements
2. Identify the goals of ALL of your programs, with attention to the goals of your programs for linguistically and culturally diverse populations.
3. Identify gaps between what is mandated and what is needed to provide evidence of student performance and development relative to local goals (*e.g., lack of common assessments of students' writing development over time*)

Identifying appropriate assessments for your system...

4. Identify common assessments that yield evidence of students' learning & achievement in specific learning areas (*i.e., common assessments of student narrative, expository, and persuasive writing that are appropriate for ALL students, particularly ELLs, including those with disabilities*).
5. Articulate assessment plan to all stakeholders and provide necessary professional development
6. Develop a timetable to guide data collection and analysis
7. Develop or adopt a data management system (*e.g., pivotal portfolio, user-friendly database that is aligned with/incorporated into district database*)
8. Collect and analyze data
9. Use data to drive instruction, program and professional development, policy, and advocacy.

What Data Do You Collect, and How Do You Use It? ¹

Complete the following survey. Use the scale (from 1-3) in the last column to rate how you use different types of data. 1 = Not at all; 2 = Somewhat; 3 = Fully

Types of Assessment Data in L1 and L2	Purposes for Assessment	Decisions	Levels of Decision-making	Data Collected	The extent that these data are used to make decisions
Standardized Measures (e.g., language proficiency tests, academic achievement tests in L1 and L2)	Summative <ul style="list-style-type: none"> • Meet accountability requirements • Place and reclassify ELLs • Improve program services 	<ul style="list-style-type: none"> • Student subgroup progress and attainment of standards • Eligibility and extent of support services 	State		1 2 3
			District		1 2 3
			Program		1 2 3
			Classroom		1 2 3
Standard or Common Measures (e.g., district writing prompts and rubrics; observation with common checklists; commonly made performance-based assessments)	Summative and Formative <ul style="list-style-type: none"> • Monitor student growth and achievement • Inform instruction • Guide program and professional development • Determine program impact 	<ul style="list-style-type: none"> • Reconfiguration of language allocation • Curriculum impact and adjustment • Regrouping of students • Effectiveness of unit or lesson • Refinement of unit or lesson 	State		1 2 3
			District		1 2 3
			Program		1 2 3
			Classroom		1 2 3
Idiosyncratic or Classroom Measures (e.g., surveys, anecdotal notes, student self-assessment)	Formative <ul style="list-style-type: none"> • Determine extent of language support • Document individual behaviors • Regulate and monitor learning 	<ul style="list-style-type: none"> • Grouping of students • Differentiation of instruction • Diagnosis of student strengths and needs • Grading • Effectiveness of unit or lesson • Refinement of unit or lesson 	Classroom		1 2 3
			Program		1 2 3
					1 2 3
					1 2 3

¹ Adapted from Gottlieb & Nguyen, 2007, Worksheet 17, pp. 219-220

Potential Uses of Assessment Information for Decision-making

- Individual student learning plans
- Classroom instructional decisions and RTI
- Focus for staff development
- School Improvement Plan
- Program/District Improvement Plan
- Public reporting and program advocacy

Illustrative Examples

1. Using common assessment data in the pivotal portfolio to foster collaboration among ESL/bilingual and mainstream teachers

- English language proficiency data (e.g., ACCESS for ELLs) and (WIDA/TESOL) can-do descriptors → Identification of concrete strategies that mainstream teacher can use to differentiate instruction and assessment for ELLs according to English language proficiency level. ESL/bilingual teacher may demonstrate how to implement strategies.
- Evidence of student writing with district rubrics → Identification of concrete strategies to improve writing instruction and group students according to strengths and needs.

2. Using common assessment data in pivotal portfolios to plan professional development

- A) English language proficiency assessment data show that ELLs are making greater gains in listening and speaking than in reading and writing and that many ELLs get stuck around Level 3 or 4 in reading and writing;
- B) General education reading assessment data show that many English speakers are not at grade level
- C) Multiple data points suggest need for reading and writing across the curriculum professional development for ALL teachers, with attention to academic language and literacy needs of diverse populations.

3. Using common assessment data in pivotal portfolio to inform teacher evaluations.

- When we have valid and reliable evidence of student growth relative to standards, goals, and benchmarks, we can use that to inform teacher evaluation.

We educators have choices in the ways that we respond to accountability requirements...

- We can react to enormous pressure of NCLB paying attention only to mandated standardized test data
- OR--
- We can take action by developing an authentic accountability system for our schools, one that is “*reciprocal and intelligent*” rather than restrictive and punitive, one that allows us to use that data to improve our programs, practices, and policies so that ALL children, including culturally and linguistically diverse students and students with disabilities, can achieve.

References

- Crawford, J. (2006). What does a valid and reliable accountability system for English language learners need to include? In Hamayan, E. & Freeman, R, *English Language Learners at School: A Guide for Administrators*. pp. 6-10. Philadelphia: Caslon.
- DuFour, R. & Eaker, R. (1998). *Professional Learning Communities: Best Practices for Enhancing Student Achievement*. Bloomington, IN: National Education Service.
- Fairbairn, S. & Jones-Vo, S. (2010). *Differentiating Instruction and Assessment for English Language Learners: A Guide for K-12 Teachers*. Philadelphia: Caslon.
- Gottlieb, M. (2003). *Large-Scale Assessment of English Language Learners*. Alexandria, VA: TESOL.
- Gottlieb, M. & Nguyen, D. (2007). *Assessment and Accountability in Language Education Programs: A Guide for Administrators and Teachers*. Philadelphia: Caslon.
- Hamayan, E. & Freeman, R. (2006). *English Language Learners at School: A Guide for Administrators*. Philadelphia: Caslon.
- Wright, W.E. (2010). *Foundations for Teaching English Language Learners: Research, Theory, Policy, and Practice*. Philadelphia: Caslon.