

2014

Accommodations and English Learners: Inconsistencies in Policies and Practice

Florin M. Mihai

N. Eleni Pappamihel



Part of the Bilingual, Multilingual, and Multicultural Education Commons, Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, and the Teacher Education and Professional Development Commons

Find similar works at: <https://stars.library.ucf.edu/tapestry>

University of Central Florida Libraries <http://library.ucf.edu>

This Article is brought to you for free and open access by STARS. It has been accepted for inclusion in TAPESTRY by an authorized editor of STARS. For more information, please contact STARS@ucf.edu.

Recommended Citation

Mihai, Florin M. and Pappamihel, N. Eleni (2014) "Accommodations and English Learners: Inconsistencies in Policies and Practice," *TAPESTRY*: Vol. 6 : Iss. 2 , Article 2.
Available at: <https://stars.library.ucf.edu/tapestry/vol6/iss2/2>

Accommodations and English Learners: Inconsistencies in Policies and Practice

Florin M. Mihai, University of Central Florida

N. Eleni Pappamihel, University of North Carolina at Wilmington

Abstract

This article outlines the results of a survey of all the large-scale testing accommodations used with English learners across the United States, addressing the increasing challenge of integrating this population into a state's accountability system in a valid and reliable manner. This examination of accommodation policies used in standardized content area assessments reveals that there is little consistency across states with regard to testing accommodations for ELs in providing accommodations when ELs are tested in academic content areas.

In today's educational environment in the United States, it is virtually impossible to find a district or school that is not held accountable by some sort of large scale or standardized assessment. In fact, as more and more federal funds are tied to the results of such assessment programs, we see this trend in accountability becoming more and more prevalent. The No Child Left Behind [NCLB] (2002) and the Common Core State Standards Initiative (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) require that all students demonstrate a high level of proficiency in mathematics and English language arts/literacy (ELA). This level of proficiency in the two content areas in presents quite a challenge for all students in schools, but even more so for those students whose native language is not English. While they are struggling to learn English and content simultaneously, they must also meet the same high academic performance expectations as native English speakers.

The number of non-native English speakers in our schools has grown rapidly, especially over the past two decades. Data from the National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs (NCELA) show that the number of English learners (ELs) in our schools continues to grow at a steady rate. In fact, between the 1994-1995 and 2009-2010 school years, while the growth of all students at the K-12 levels has remained steady between 2 and 4 %, the population of ELs rose by over 60% (NCELA, 2011).

In the past, states were not required to include ELs in their accountability plans, often leaving this population in a state of limbo with regard to their actual academic and linguistic progress. However, with the passage of Goals 2000 (H.R. 1804) standards and assessments had to apply to all students, including ELs. Moreover, in the year 2000-2001, each state was required to have an assessment tool or instrument that included ELs and ensured that these students make 'adequate yearly progress' (Menken, 2000). As previously mentioned, these requirements have been reinforced and extended through the reauthorization of the Elementary and Secondary Education Act or No Child Left Behind and the Common Core State Standards Initiative. States

<http://tapestry.usf.edu/journal>

are now held accountable for progress in content areas as well as in English language proficiency standards and must disaggregate data for ELs to show their academic achievement.

While increased accountability for ELs contributes to ensure academic progress for this group, Boyd-Batstone (2013) makes a very salient point when stating that there is no single assessment to address the needs of all students. There is great controversy as to whether or not results from large-scale tests are valid and reliable (Abella, Urritia, & Shneyderman, 2006; Wise, Hauser, Mitchell & Feuer, 1999; Wright, 2010). In order to increase the validity of these results, all states have instituted a variety of accommodations for ELs when taking standardized large-scale assessments, and researchers have begun to collect data on the effectiveness of a variety of such accommodations. However, the lack of a standard in the manner in which accommodations are implemented can be problematic and cast doubt upon the effective incorporation of research findings into accountability systems. Not only are the accommodations provided to ELs on large scale assessments not consistently applied across states, but there is little empirical evidence supporting their effectiveness under specific implementation circumstances.

What is an Accommodation?

According to Butler and Stevens (1997) an accommodation for ELs on large-scale assessments, “refers to the support provided students for a given testing event, either through modification of the test itself or through modification of the testing procedure, to help students access the content in English and better demonstrate what they know “ (p. 5). Modifications of the test may include bilingual translations, glossaries (either in English or the student’s native language), and/or simplified text. Modification of the test procedures can include flexible time limits, flexible setting, and/or the reading and explanation of directions.

To this broad definition, Abedi and his colleagues add a differentiation that targets the effectiveness of accommodations. Abedi, Hofstetter, and Lord (2004) state that, in order to be effective, accommodations should level the playing field for the ELs by helping them overcome the language barrier and thus assessing what they know in terms of content as opposed to language. In other words, a true accommodation does not benefit ELs over native English speakers. An example of this would be using a specialized glossary as opposed to a standard dictionary. In English there are several definitions for the word “bear”. In the question, “How much weight can the fixture bear?” if an EL looks in the dictionary, there are over 10 separate definitions for this word in its verb form alone. Within the appropriate usage, which is ‘to support’, there are six variations on that theme. Whereas a native speaker of the language would likely not need any help in defining that word in that context, an EL could easily be overwhelmed by these numerous possibilities. A specialized glossary would supply only the definition appropriate to that context. In this case, because only the EL is aided by that accommodation, nothing is withheld from the native speaker and the validity of the test is preserved.

Effectiveness of Accommodations

The most important question to be answered is whether or not these accommodations are effective in providing valid and reliable data, even within their own states. As part of the data collection for this article consisting of examining EL accommodation policies across states, the authors wrote to several states with large EL populations asking whether or not any research had been done to determine if the accommodations offered had been demonstrated to be effective; we received no response from any state we contacted. In fact, Abedi, Hofstetter, and Lord (2004) write that few decisions concerning accommodations are made based on empirical evidence regarding their validity.

In light of the lack of information from the states themselves, we turn to research done in other contexts. According to Abedi and his colleagues, the most successful accommodation is language simplification. Districts making use of this accommodation would rewrite tests in simplified language to reduce the linguistic burden of the assessment while maintaining a high standard of content knowledge (Abedi, Lord, & Plummer, 1997). In fact, of four common accommodations (extra time, glossary, linguistically modified items, and glossary plus extra time), Abedi and his colleagues found that only simplification of text reduced the performance gap that exists between ELs and native English speaking students (Abedi, Lord, Hofstetter & Baker, 2000; Abedi, Lord & Hofstetter, 1998). After looking at EL performance, the researchers discovered that translating items into the EL's native language did not appear to be an effective accommodation strategy for those students who had studied the subject in a classroom where English was the primary language spoken.

Another crucial finding was that the performance gap between English learners and other students was narrowed by changing the language of the test, i.e. to reduce the use of low-frequency vocabulary and complex language structures that are secondary to content knowledge being assessed. More importantly, this accommodation strategy was not only effective but also valid, since it did not appear to affect the performance of English-proficient students. This result has been supported in other studies where it has been found not to pose an unfair advantage for ELs and to have little or no impact on the test scores of native English speakers (Rivera & Stansfield, 2004).

Hence, the main recommendation from these studies is that there is no 'one size fits all' solution but that language simplification seems to provide the most useful method of reducing the achievement gap. However, this recommendation is not without controversy. Zehr (2006) has cited several other researchers in the field who do not necessarily disagree with Abedi but rather point out that the research base is limited. Zehr cites Charlene Rivera of the Center for Equity and Excellence in Education who believes that too little research has been done to conclusively support one recommendation over another. Margo Gottlieb of the Illinois Research Center concurs with this position, and states that further study is warranted.

Given the inconsistency in opinions on this topic, it is unlikely that any state can be sure that the test results for ELs are both valid and reliable. NCLB requires that "English language learners must be tested in a 'valid and reliable' manner. It says that states must provide accommodations for such students and, 'to the extent possible,' test them in a 'language and form most likely to yield accurate data on what students know and do.'" (Zehr, 2006, p. 4).

Several states started implementation of tests in the native language of the ELs, but, as noted previously, Abedi and others have not found it to be the most effective accommodation. Other states allowed the use of dictionaries, employed scribes and translators, or simply gave their ELs more time to take the test. Again, the capacity of such accommodations to yield accurate data has not been consistently proven, yet, they affect test outcomes and resulting educational decisions for over 5 million English learners enrolled in the US public school system.

Moreover, with so many states using such a variety of accommodations, it is difficult to compare test scores across the US. Even when an accommodation falls into a single overall category, often the implementation of that accommodation is dramatically different in different states. For example, a common type of accommodation, translation, is quite often employed differently. Some states offer bilingual tests (where both languages are presented to the student) whereas other states offer monolingual translations (the student only sees one language) or clarification in the child's native language when requested. Other states offer oral translations of test items. Yet, we do not know which of these types of translation is valid and reliable, and it would be irresponsible to regard them equally without empirical confirmation.

Inconsistencies in Accommodations

Before the passage of NCLB (2002), Butler and Stevens (1997) identified two broad categories of testing accommodations. These consisted of modifications of the test, which included text changes, visual supports, glossaries, translations, and other linguistic supports, and modifications of test procedures, which included breaks, extra time, small group administration and reading aloud among others. However, this distinction is relatively broad, and we feel is no longer as useful since the current legislation is requiring the inclusion of ELs in statewide assessment programs. In the past, many states offered accommodations for ELs. However, some did not have specific guidelines and EL accommodations were lumped together with accommodations provided for students with disabilities. What we have found in our research is that in spite of the assessment requirements of NCLB and CCSS, some of this ambiguity still remains. However, more and more states now distinguish between EL and Special Education accommodations as opposed to applying Special Education accommodations to ELs without discrimination.

Koenig and Bachman (2004) addressed this somewhat broad categorization of accommodations by organizing them into four categories instead of two: presentation, response, time and scheduling, and setting. Presentation accommodations involve, for instance, translation of the directions and translation of the test into the student's native language. Response accommodations allow ELs to respond in the native language or in both the native language and English. Time and scheduling accommodations give ELs extra breaks or extended testing sessions over multiple days. Setting accommodations include small-group, separate-room, and individual administration of the test.

Nevertheless, even a four-category system is still relatively broad. It is beneficial to distinguish the types of accommodations even further in an effort to better serve ELs by eliminating any source of ambiguity and confusion.

Expanding on the four-category system (Koenig & Bachman, 2014), we propose that testing accommodations for ELs be divided into six categories including modifications to time (schedule changes, additional time), testing environment (preferential seating, private setting), testing aids (dictionaries, computers, etc.), testing directions (in native language, read aloud), test presentation (bilingual version, audio/video recorded) and test response type (in native language, point to responses in English) (Mihai, 2010).

The most commonly used accommodations fall into several of our 6 categories. These include additional time (39 states), use of dictionaries (38 states), individual/small group administration (34 states), flexible scheduling (28 states), and reading aloud of the test in English (25 states). These accommodation categories are most often the descriptors seen on state testing sites and, on the surface, seem to be quite consistent. However, when we examine the accommodations in more detail, we find that the implementation of the accommodations is quite varied and could impact the validity and comparison of the test scores in many ways. In order to examine how each state operationalizes each of the six categories, we analyze the components of each accommodation.

Possibly the most consistent accommodation is flexible scheduling because of the ease with which this accommodation can be implemented. However, even within this relatively straightforward accommodation, there is variety among states. Some states such as Alabama, California, and Wisconsin, allow ELs to take extra breaks during testing, but others such as Florida and New York do not. This inconsistency could impact validity in terms of test fatigue that could be compounded in students who are forced to complete tests in a given day without being allowed to take additional breaks. Even though both groups are seen as having the same accommodation, students in Florida and New York are obviously not testing under the same conditions.

In terms of individual/small administration, the inconsistencies become more pronounced. Two states (Michigan and Wisconsin) allow ELs to move about the classroom during testing. Idaho, Kentucky, Michigan and New Jersey only allow small group testing; yet, California and Hawai'i allow individual testing. Other states (Georgia, Texas) allow either individual or small testing, depending on the needs of the EL or school. Twenty-one states allow 'flexible setting', which means that ELs are allowed to take their tests outside the regular classroom. However, other states include a category known as 'flexible seating' that keeps ELs in the regular classroom for testing but allows them to move to a different seat in the classroom. These two different iterations of a similar accommodation can have a significant impact on validity.

There are 20 possible accommodations (Table 1) involving testing aids, which can be broadly divided into two subcategories. The first subcategory includes accommodations striving to reduce the linguistic burden, whereas the second one is represented by accommodations aimed at reducing distractions and increasing focus. In the first sub-category, the most wide-spread type of accommodation is allowing ELs to use dictionaries when taking the test. However, we find critical differences among states. For example, some states allow bilingual dictionaries, while others (Arkansas and Massachusetts) allow such dictionaries but without definitions or pictures. Some states (Wisconsin and Mississippi) allow spelling dictionaries, while others categorically forbid them. To emphasize this inconsistency, Texas mentions the use of English language

dictionary as an accommodation but does not recommend a specific type of dictionary. Given that dictionaries are so varied in type (print, electronic), language (bilingual, monolingual), usage (translation, definition), this level of generalization is unacceptable. Each dictionary-based accommodation may have a different impact on reducing the linguistic burden and thus affecting the validity of test scores.

Accommodation	Number of States
Spelling dictionaries	2
English language dictionaries	1
General dictionaries	2
Translation/bilingual dictionaries	21
Bilingual dictionaries without pictures/definitions	13
Electronic/translation dictionaries	6
Electronic/print dictionary in English or native language	1
English/Native language glossaries with no definitions	1
Translation glossaries	2
Student-made glossaries	1
General glossaries	3
Word lists	1
English/Native language word lists with no definitions	1
Bilingual vocabulary lists	2
Use typewriter/computer spell check off	4
Word processor	7
Use computers with spell check on	2
Markers	1
Noise buffers	2
Assistive technology	5

Table 1: All accommodations involving testing aids

Another accommodation in this subcategory is the use of glossaries and word lists, from translation glossaries to bilingual vocabulary word lists. Some types of glossaries are student made. Some include definitions, and others include the ELs' native language. While previously mentioned research has indicated that glossaries can be a very effective accommodation, the lack of consistency in implementation of this accommodation questions that effectiveness.

There are twenty accommodations that affect test directions. All are directed towards the reduction of the linguistic burden for ELs. The most commonly used accommodations in this category address both English and the students' native language. The fundamental difference in accommodation philosophy is whether or not the accommodation is provided in the ELs' native

language or only in English. A common statement found in testing regulations is, “Directions _____ in English.” or “Directions _____ in native language.” These blanks can be filled in with a variety of words including, ‘read’, ‘repeated’, ‘clarified’, ‘explained’, ‘paraphrased’, ‘simplified’, ‘rephrased’, or in the case of a native language accommodation, ‘translated’. However, there is a great deal of ambiguity in exactly what these words actually mean on test day that can affect validity. Even if test administrators in Idaho know what it means to provide a test direction clarification, test administrators in Montana may view test direction clarification differently. The same type of variation applies to test presentation in which ELs across the country may have access to completely different versions of a similar accommodation. This can vary from a video presentation of the test to a native language read-aloud of the test. Both would be viewed as native language presentations of the test, but in reality could be seen by the test-taker as very different.

Test response accommodations are also varied and are not as simple as whether or not students are able to use their native language or not. In some cases, ELs are allowed to dictate responses to scribes in their native language (Nebraska), but in other cases their responses are transcribed to English at a later date (Maryland, Wisconsin, and Arizona). This simple difference could potentially introduce several kinds of threat to validity to test scores including translation issues among others.

Implications and Conclusions

This examination of accommodation policies reveals that there is little consistency across states with regard to testing accommodations for ELs. This inconsistency could be related to the lack of information regarding the effectiveness of testing accommodations for ELs. While most of the accommodations for ELs are modeled after testing accommodations made for students with disabilities (Zehr, 2006), which has a much longer history, it is clear that ELs and students with disabilities face dramatically different issues when taking large-scale assessments.

States obviously use accommodations to level the playing field for ELs in high stakes testing programs to increase the validity of test scores. Yet, as we have discussed, the actual validity that these accommodations add to test scores is in serious question. Are the scores for ELs valid and reliable? At this point in time, we do not know this within a reasonable doubt.

Currently, states are required to make valid instruments to measure content growth, and in the case of ELs, language proficiency. While many states are making progress in measuring English proficiency, the United States Department of Education is taking the lead in helping states develop guiding standards so that we can all be consistent in using English proficiency standards. We need a similar consistency in providing accommodations when we test ELs in academic content areas. Without this careful consideration, we face the possibility of once again invalidating the test results of ELs and blaming the victim for the failures of the system.

References

- Abedi, J., Hofstetter, C.H., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Education Research, 74*(1), 1-28.
- Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000) *Impact of accommodation strategies on English language learners' test performance. Educational Measurement: Issues and Practice, 19*(3), 16-26.
- Abedi, J., Lord C., & Plummer, J. R. (1997). *Final report of language background as a variable in NAEP mathematics performance*. Los Angeles: Center for the Study of Evaluation, CSE Technical Report # 429.
- Abedi, J., Lord, C., & Hofstetter, C. (1998). *Impact of selected background variables on students' NAEP math performance*. Los Angeles: UCLA Center for the Study of Evaluation/ National Center for Research on Evaluation, Standards, and Student Testing.
- Abella, R. Urritia, J. & Shneyderman, A. (2005). An Examination of the Validity of English-Language Achievement Test Scores in an English Language Learner Population. *The Bilingual Research Journal, 29*(1), 127-144.
- Boyd-Batstone, P. (2013). *Helping English Language Learners meet the Common Core: Assessment and instruction*. New York, NY: EyeOnEducation/Routledge.
- Butler, F. A., & Stevens, R. (1997). Accommodation strategies for English language learners on large-scale assessments: Student characteristics and other considerations (CSE Tech. Rep. No. 448). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Goals 2000 (1994). <http://www.ed.gov/legislation/GOALS2000/TheAct/index.html>. (accessed August 15, 2006).
- Kindler, A. (2002). Survey of the states' Limited English Proficient students and available educational programs and services 2000-2001 summary report. <http://www.ncela.gwu.edu/policy/states/reports/seareports/0001/sea0001.pdf> (accessed September 21, 2006)
- Koenig, J. A., & Bachman, L. F. (Eds.). (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessments*. Washington, DC: National Academies Press
- Menken, K. (2000). *What are the critical issues in wide-scale assessment of English language learners?* <http://www.ncela.gwu.edu/pubs/issuebriefs/ib6.htm> (accessed September 21, 2006).
- Mihai, F.M. (2010). *Assessing English language learners in the content areas: A research into practice guide for educators*. Ann Arbor, MI: The University of Michigan Press.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: Authors.
- No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).
- NCELA (2011). *The growing number of limited English proficient students 2009/2010*. http://www.ncela.us/files/uploads/9/growing_EL_0910.pdf. (accessed April 29, 2014)
- Rivera, C. & Stansfield, C.W. (2004). The effect of linguistic simplification of science test items on score comparability, *Educational Assessment, 9*, 79-105.
- Wise, L.K., Hauser R.M, K.J. Mitchell, K.J and Feuer, F.J. (1999). *Evaluation of the voluntary national test: Phase I*. Washington, DC: National Academy Press.

Wright, W.E. (2010). *Foundations for teaching English language learners: Research, theory, policy and practice*. Philadelphia, PA: Caslon.

Zehr, M.A. (2006). Scholars seek best ways to assess English-learners. *Education Week*, 25(18), 10-10.

About the Authors

Florin M. Mihai, Ph.D., is an Associate Professor in the Department of Modern Languages and Literatures at the University of Central Florida. His research interests include: language and content-area assessment for English learners, grammar instruction, pre- and in-service teacher education, and curriculum development in global contexts.

N.Eleni Pappamihel, Ph.D., is an Associate Professor in the Department of Instructional Technology, Foundations, and Secondary Education at the University of North Carolina at Wilmington. Her main area of research focuses on improving academic outcomes for English language learners in the mainstream classroom and includes work in co-teaching and digital storytelling as an instructional strategy.