Validity issues in designing accommodations for English language learners

Jamal Abedi

Introduction

Different accommodations are used in the assessment of students who are English language learners (ELLs). However, many of these accommodations were originally created and used for students with disabilities (see Oller, this volume); therefore, the utility of such accommodations for ELL students is questionable. To help reduce the impact of construct-irrelevant sources in the assessment of ELL students, accommodations should be used that (1) are effective in making assessments more accessible for ELL students, (2) provide assessment outcomes that are valid, i.e., comparable with those of non-ELLs, and (3) are sensitive to student background. These three characteristics are vital in determining the validity of accommodated outcomes. First, if an accommodation is not helpful in reducing the performance gap which is partly due to the impact of construct-irrelevant sources, then the accommodation is not doing what it is intended to do. For example, if an accommodation does not help ELL students with understanding the complex linguistic structure of the assessment, which is unrelated to the content being measured, then that accommodation may not be relevant. Second, if an accommodation alters the construct being measured, then the validity of the accommodated assessment could be at risk. An accommodation that provides unfair advantage to the recipients may impact the measurement of the construct. Third, if an accommodation does not address each individual student’s educational needs, then the outcome of the accommodated assessment may not present a good picture of what the student truly knows and is able to do. For instance, providing a native language assessment to an ELL student who has been predominantly instructed in English may not produce a desirable outcome. In this chapter, we introduce the concept of accommodation for ELL students, discuss major issues concerning accommodations, and elaborate on the limitations of the currently used accommodations for ELL students. The chapter also discusses the methodology for conducting studies to examine the effectiveness and validity of accommodated assessments and provides recommendations on best practices of accommodations.

Historical perspectives

In order to provide fair assessment for every child in the United States, both federal (No Child Left Behind Act of 2001, P.L. 107–10) and state legislation require the inclusion of all students,
including ELLs, in state and national assessments. There are, however, major technical issues in including ELL students in large-scale assessments. The limited English proficiency of these students may create unequal opportunities for them to benefit from the teacher’s instructions and fully demonstrate their knowledge and abilities (Abedi and Herman, 2010). In an attempt to level the playing field, current legislation requires that these students receive accommodations—changes in the test process, in the test itself, or in the test response format. The objective of accommodations is to provide a fair opportunity for English language learners to demonstrate what they know and what they are able to do, without giving them an advantage over students who do not receive accommodations.

Many states currently use accommodations for ELL students (Abedi, 2007; Abedi et al., 2004; Francis et al., 2006; Rivera et al., 2000; Willner et al., 2008; Sireci et al., 2003; Wolf et al., 2008). The widespread use of accommodations, however, raises a number of issues and questions. Does the use of accommodations provide more valid inferences about ELL students’ content knowledge? Which students should be eligible for receiving accommodation(s) and what criteria should be used to decide their eligibility? What type of accommodation should be used? Are some accommodations more effective than others for individual students with different background characteristics? Do accommodations provide an unfair advantage to the recipients? Is it meaningful to aggregate accommodated and non-accommodated scores? When we look for answers to these questions throughout studies of content area assessments, we are confronted with a striking lack of empirical research.

Critical issues and topics

The questions and concerns that were presented above can be organized into the following four questions, all of which need to be fully addressed before one chooses to use accommodations in the assessments of ELL students.

1. How effective are the accommodations in making assessments more accessible for ELL students? (Effectiveness)
2. Are the accommodated assessments as valid as the non-accommodated assessments? (Validity)
3. Do accommodations work the same for all ELL students, despite their different backgrounds and their differences in the levels of English proficiency? (Differential impact)
4. How relevant are the accommodations that are used for ELL students to their specific needs? (Relevance)

To shed light on these questions, one must examine these issues using studies that are based on a sound research methodology. Among the necessary conditions for such experiments are random assignments of ELL and non-ELL students to different accommodation conditions.

The most commonly used accommodations for ELL students within the United States are extended time (42 of the 48 states); use of a glossary (26 states), an English dictionary (33 states) or bilingual dictionary (22 states); and linguistically simplified test items (12 states) (Rivera, 2003). Willner et al. (2008) distinguished between accommodations that are relevant for ELLs and those that may be irrelevant for these students. Research results are available only for a few forms of accommodations, while others have yet to be examined. For example, research has shown that providing an English dictionary (Abedi et al., 2003a) and extra time (Abedi et al., 2004; Hafner, 2001; Thurlow, 2001) affects the performance of all students, including those for whom the accommodations are not intended (see also, Maihoff, 2002; Thurlow and Liu, 2001). To aggregate the results of accommodated and non-accommodated assessments, one must have enough evidence on the validity of accommodations.
Results of research on the assessment and accommodations of ELL students have indicated that some accommodations are not effective in making assessments more accessible to ELL students while others have been shown to have a more positive impact on the assessment outcomes for ELL students. Among them are linguistic modification of assessments, English dictionary/glossary and bilingual dictionary/glossary, and testing in the student’s native language. In the next section of this chapter, we will present a summary of research on some of the most commonly used accommodations in the nation and will then briefly present methodological issues for examining major characteristics of these accommodations based on the four research questions presented above.

Current contributions and research

As indicated earlier, accommodations are meant to assist ELL students in the area where they need the most help, i.e., language, in order to “level the playing field” for these students. Accommodations are strategies intended to reduce threats to validity of test scores. Students’ performance on content-based assessments, such as mathematics and science, can be confounded by language, which is considered to be irrelevant to the construct being measured. In other words, the test may unintentionally measure students’ language abilities, rather than their content knowledge in areas such as math or science. Accommodations can help ELL students demonstrate their content knowledge by reducing the confounding of language with the content knowledge. However, accommodations are not intended to give ELL students an unfair advantage over students not receiving accommodations. Therefore, research must be conducted to determine not only if an accommodation helps ELL students score better on an assessment than without it, but also that the accommodation does not alter the construct being measured. Below is a brief presentation of research findings for the following most commonly used accommodations for ELLs:

- extended time
- English dictionary
- English glossary
- bilingual dictionary or glossary
- customized dictionary
- native language testing
- linguistically modified test
- computer testing.

Extended time

This accommodation provides students with additional (or extended) time to complete an assessment. This is one of the most commonly used accommodations for ELL students and students in different categories of disabilities (Rivera et al., 2000) possibly due to its ease of administration relative to other accommodations. Some studies found extended time to be an effective accommodation for ELLs and students with disabilities (Abedi et al., 2000; Hafner, 2001; Thurlow, 2001; Chiu and Pearson, 1999).

However, other studies did not show extended time to be effective for ELLs and students with disabilities (Munger and Loyd, 1991; Fuchs et al., 2000; Marquart, 2000). A major issue is that extended time may affect the performance of non-ELL students positively as well, therefore making the validity of an assessment using this accommodation suspect (Abedi et al., 2000, 2004; Wolf et al., 2008).
English dictionary

Providing a published English dictionary for students to use during a test is another commonly used accommodation for ELL students (Abedi et al., 2000; Sireci et al., 2003; Willner et al., 2008). However, a dictionary must be provided in combination with extended time in order to avoid the problem of information overload. Some studies have found that the dictionary and extended time affect the performance of all students (Abedi et al., 2000; Hafner, 2001; Thurlow, 2001; Maihoff, 2002; Thurlow and Liu, 2001; Wolf et al., 2008), casting doubt on the validity of the assessment outcomes under this accommodation.

Furthermore, by having access to the definitions of content-related terms, recipients of a dictionary may be advantaged over those who do not have access to dictionaries, which may compromise the validity of assessment. Consequently, the results of accommodated and non-accommodated assessment may not be aggregated. Providing a dictionary is also difficult and burdensome for test administrators, which makes it less feasible to implement (Abedi et al., 2001).

English glossary

An English language glossary can be provided as an alternative to a dictionary. In this case, only terminology that appears in the test appears in the glossary, making a glossary less burdensome than providing a dictionary. In some studies the English glossary with extended time raised the performance of both ELL and non-ELL students. For example, the results of a study by Abedi et al. (1998, 2000) showed ELL students’ performances increased by 13% when they were given a glossary with extended time. While this seemed promising, the non-ELL students’ performances increased by 16% (Abedi et al., 2004), indicating a change in construct from the non-accommodated test version. That is, rather than decreasing performance gap between ELLs and non-ELLS, this accommodation actually increased the performance gap between the two groups. Thus, the outcomes of assessments under this accommodation may not be aggregated with the non-accommodated outcomes.

Bilingual dictionary/glossary

Bilingual dictionaries and glossaries are also among widely used accommodations for ELL students. As with an English dictionary, by having access to definitions of content-related terms, recipients of a published bilingual dictionary may be advantaged over those who did not have access to one (Abedi et al., 2000, 2003b) Another major limitation with a bilingual dictionary is the content equity issue. Different published bilingual dictionaries present a substantial range of content coverage. Furthermore, students who speak another language at home may not be literate or fully literate in their home language, and would not find a bilingual dictionary helpful. They may also be unaccustomed to using one in the classroom, making it unfamiliar, and therefore not useful.

Customized dictionary

In one study, a customized English dictionary was introduced and used as a more valid alternative to English/ bilingual dictionaries (Abedi et al., 2001). The customized dictionary was a literal cut-and-paste of the actual dictionary entries, which only included terms that were (1) in the test and (2) non-content related. Results of studies suggested that it is a highly effective and valid accommodation for ELL students.
Native language testing

Translating tests into a student’s native language is an accommodation used by many states across the country (Abedi et al., 2001; Rivera et al., 2000). However, translating a test can make the instrument easier or harder in another language, and some cultural phrases and idioms can be difficult to translate (Hambleton, 2001). Furthermore, students may be only proficient in speaking their home language, not in reading it, particularly if the students have not received formal instruction in the home language. Translating assessment tools into a student’s native language may not produce desirable results if the language of instruction and assessment are not aligned (Abedi et al., 2000). Given the large number of languages spoken in public schools, translating tests also raises many feasibility issues due to the resources needed.

Linguistically modified test

Some accommodations help ELL students with their English language needs without compromising the validity of assessment. Studies suggest that a linguistically modified version of the test items is an effective and valid accommodation for ELL students (Abedi et al., 2000; Maihoff, 2002; Rivera and Stansfield, 2001). We will elaborate more on this form of accommodation for two reasons. First, as indicated above, a linguistically modified version of a test is a good example of an effective, valid, and relevant accommodation for ELL students, and, second, the impact of linguistic factors in the assessment of ELL students is clearly being demonstrated by this accommodation.

Linguistic modification involves modifying the linguistic features on existing assessments of any item that may include unnecessary complexity. For example, something as simple as switching from passive voice to active voice is considered linguistic modification. Some research refers to this as linguistic simplification, but it is not necessarily simplifying a test, but rather reducing linguistic complexity that is not related to the construct.

To provide a more contextual framework for this accommodation, we start with the concept of language as a source of construct-irrelevant variance for content areas in which language is not the primary construct being measured. Research has identified 14 major categories of linguistic features that appear to contribute to the difficulty of comprehending text (Abedi, 2006, 2007, 2010; Abedi et al., 1997; Shaftel et al., 2006). These features may slow the reader down, increase the likelihood of misinterpretation, or add to the reader’s cognitive load, thus interfering with concurrent tasks. Indexes used to measure language difficulty include unfamiliar (or less commonly used) vocabulary, complex grammatical structures, and discourse that include extra material, abstractions, and passive voice (Abedi et al., 1997).

In the linguistic modification approach, sources of unnecessary linguistic complexity, such as those mentioned above, are identified. A team of content, linguistic, and measurement experts determine which linguistic features, among the 14 major categories that were identified in research (Abedi, 2010), are present in the text and how to reduce their impacts. First, each item in the test is examined for the presence and intensity of each linguistic feature by rating these features in the item. A five-point Likert-type rating rubric (with 1 indicating little or no complexity of the item on that feature and 5 suggesting that the item was seriously impacted by the feature) is used to generate ratings on all items. Items that were rated high (rating of 3 or higher) can then be marked for either deletion or revision. The same group of experts who identified these features would then guide the revision process, changing or eliminating items when necessary. The goal would be for all items to have a rating of 2 or below on the test.

This accommodation may have great potential to be effective and help provide valid assessment outcomes for ELL students. It appears to be quite relevant for ELL students, since it provides the
type of assistance (linguistic assistance) that ELL students need, is based on consistent results from experimentally controlled experiments, and does not alter the construct being measured (Abedi, 2010).

**Computer testing**

Research findings suggest that computer testing is an effective and valid accommodation for ELL students. In a study by Abedi (Abedi, 2009; Abedi et al., 2003), ELL students showed higher levels of motivation on the assessment when it was administered by computer. The study used a computer to implement a customized English “pop-up” glossary, which was also found to be effective and valid for ELL students. The “pop-up” glossary aspect of the system was set up to only gloss terms that were unrelated to the content being measured (mathematics in this study). The system kept track of how often the glossary terms were utilized and how much time each student spent on the glossary terms. In a modified second version of this study (currently being conducted), several additional accommodation features have been added to the system. These new features include variation in font size, read aloud test items and adaptive testing based on the students’ level of English proficiency, which is tested at the start of the session.

**Main research methods**

The summary of literature presented above clearly shows mixed results on the effectiveness and validity of many forms of the accommodations currently used for ELL students. There might be several issues with the design of these studies that could lead to such mixed findings. First and foremost are the issues regarding samples of students for these studies. If the sample size is small or the sample is not representative of the population in question, then inconsistent results could be obtained. The inconsistencies in the outcomes of the studies described above may also be due to the impact of extraneous variables on the study that are considered as threats to internal and external validity of the study that may not have been properly controlled (Gay, 1981). To help understand methodological issues in the use and interpretation of accommodations, we will discuss the methodological issues concerning accommodations that were reflected in the four research questions presented earlier.

**Question 1: examining the effectiveness of accommodations for ELL students**

The concept of accommodation was developed and applied mainly in the assessment of students with disabilities, and some of the accommodations that are used for ELL students were originally designed and implemented for students with disabilities (Willner et al., 2008). Therefore, they may be irrelevant for ELLs and may not be effective at improving ELL student performance. For example, among the 73 accommodations used for ELL students across the nation (Rivera, 2003), only 11 (15%) of them were found to be relevant for ELL students (Abedi, 2006). The rest of these accommodations (62) did not seem to be relevant or did not even have face validity for ELL students.

An accommodation is effective for ELL students if it helps reduce the sources of construct-irrelevant variance and make the assessment more accessible to these students. Therefore, it is imperative to assess ELL students’ needs and the ways that accommodations could help them when selecting effective accommodations. Unlike students with disabilities, who are grouped under several different categories with different types of disabilities and different needs, ELL students share a common characteristic: they all need linguistic assistance. Therefore, accommodations that could directly address their linguistic needs would be the most relevant accommodations for this group.
Effective accommodations for ELL students should help to provide assessments that are more accessible to these students by reducing unnecessary linguistic complexity of assessments. The effectiveness of accommodations for ELL students can be examined through an analysis of performance differences between ELL students assessed under the accommodation conditions (the treatment group) and ELL students tested under the standard testing condition with no accommodations provided (the control group). This comparison should be done through a randomized experiment in which students are assigned randomly to the treatment (accommodated) and control (non-accommodated) groups. Such random assignment (assuming a large enough sample size) would control for many sources of threats to internal validity of the design for testing effectiveness of accommodations. By assigning students randomly to the treatment and control groups students’ initial differences can also be controlled.

Unfortunately, however, results of studies on the effectiveness of accommodations have illustrated that some of the accommodations that are shown to be effective may alter the construct being measured (Abedi et al., 2000; Sireci et al., 2003). Therefore, it is essential that evidence on the effectiveness of accommodations be interpreted in light of the validity evidence for the accommodations.

**Research design for examining effectiveness of accommodations for ELLs**

ELL students should be randomly assigned to either a treatment group in which they receive accommodations or to a control group in which they are tested under the standard testing condition in which no accommodation is provided. Evidence of improvement in the performance of ELL students under the accommodated condition may be an indication of the effectiveness of the accommodation used. However, as indicated above, the study of how effective each accommodation is must be conducted under experimentally controlled conditions by assigning subjects randomly to the treatment and control groups so that any possible initial differences that may affect the outcome of the study could be controlled. It is also important to control for any other sources of variations that could affect the outcome of the effectiveness study. For example, studies on accommodations are often conducted in intact classrooms. There are some expected sources of variability due to classroom characteristics (e.g., teacher background, average socioeconomic status, percentage of ELLs, percentage of students with disabilities) and school and district characteristics and resources that can be controlled statistically with a hierarchical linear modeling approach and other statistical methodologies.

**Question 2: examining the validity of accommodations for ELL students**

Accommodations are provided to ELL students to reduce construct-irrelevant variance from the assessment outcomes; thus making assessments more accessible to ELL students. A major source of construct-irrelevant variance is due to language factors in the assessments of ELL students. Unnecessary linguistic complexity in the assessment introduces an additional dimension into the assessment model, which may be unrelated to the target construct being measured. However, in the use of accommodation one must be extremely careful that the accommodations used do not alter the construct being measured. Therefore, one approach in examining the validity of accommodations is to compare the performance of non-ELL students under the accommodated and non-accommodated conditions. If the performance of non-ELL students, for whom accommodations are not intended, changes significantly under the accommodated condition, then the accommodations is believed to do more than what is intended to do. The accommodated outcomes under such situation may not be valid. Another approach in examining the validity of
accommodations is to compare the structural relationship of the assessment outcomes with a set of valid external criteria under accommodated and non-accommodated conditions.

While there is not enough evidence to judge the validity of all the accommodations used for ELL students, results of existing research suggest that some of the accommodations used for ELL students and students with disabilities may alter the construct being measured (Abedi, 2007; Bielinski et al., 2001; Meloy et al., 2000). For others, however, there is not much evidence to judge the validity of assessments under those accommodations. Such lack of evidence puts states in a very difficult position. They may have to use accommodations without having evidence on their validity and comparability.

**Research design for examining validity of accommodations for ELLs**

Both effectiveness and validity of accommodations depend, to a large extent, on the content being measured. Accommodations that may be effective and valid for one content area may not be so for other content areas. Therefore, it is important to examine the validity of accommodations for each particular content and testing situation. To provide more valid and consistent results on the validity of accommodations, an experimental study can be conducted where large samples of students are randomly assigned to the accommodated and non-accommodated conditions and sources of extraneous variables that affect validity of the outcomes are controlled to the extent possible. Then, the validity of the accommodations can be examined using two different yet complementary approaches: (1) performance difference and (2) the criterion-related validity approach.

**Performance difference**

If accommodations do not alter the construct being measured, then performance of non-ELL students—for whom accommodations are NOT intended—should remain the same under both the accommodated and non-accommodated assessments. Thus, a significant performance change under accommodated assessment for the non-ELL students may suggest that accommodations are doing more than what they are supposed to do. The null hypothesis on the validity of the accommodated assessments can be examined separately for each accommodation. To test this hypothesis, the mean scores of non-ELL students are compared across the accommodation conditions (accommodated/non-accommodated). If the null hypothesis of no performance difference between the accommodated and non-accommodated assessments is rejected, then one can conclude that the accommodations are affecting the constructs being measured and that the validity of those accommodations may have been compromised.

**Criterion-related validation approach in examining validity of accommodations**

Some researchers argue that increased performance of non-ELL students under accommodations may not be enough to suggest that the accommodated assessment outcome is invalid, especially when the ELL group benefits more from the accommodations (the interaction effect). Therefore, examining the validity of accommodated assessments using the criterion-related validity approach may provide more convincing evidence for judging the validity of accommodated assessments. It would be helpful to apply both approaches in the study of the validity of accommodations.

To assess the validity of accommodation through an external criterion approach, additional information from the participating students would be needed. Students’ state assessments scores, as well as some relevant background variables, may be used as the criterion variables. These variables can then be correlated with the accommodated assessment scores using a latent-variable modeling approach. Using a multiple-group confirmatory factor analysis approach, the structural
relationships across the accommodated and non-accommodated assessments can be compared (Abedi, 2002). Hypotheses of differences between the two groups (non-ELL accommodated and non-ELL non-accommodated students) can be examined by testing the invariance of the relationships between the accommodated assessment outcomes and the external criteria.

The following hypotheses of invariance can then be tested:

1. Factor loading of item parcels with the latent variables for math would be the same across the accommodated and non-accommodated assessments.
2. Correlations between the math test and the state assessment latent variables (math) would be the same across the accommodated and non-accommodated assessments.
3. The fit indices would be similar between the treatment and the control groups.

If these hypotheses of invariance are not rejected, then the comparability of accommodated and non-accommodated assessment outcomes can be established.

In summary, there are concerns over the validity of accommodations used in the assessment of English language learners. For many of the currently used accommodations there is not sufficient evidence on how they perform in term of validity of assessment outcomes. Without convincing evidence on the validity of accommodations used in practice, it would be difficult to aggregate the accommodated and non-accommodated assessment outcome for assessment and/or accountability purpose.

**Question 3: examining the differential impact of accommodations for ELL students**

The level of impact accommodations have on student performance may depend on a student’s academic background. That is, a single accommodation may not work for all students. As indicated earlier in this chapter, ELL students share a common need and that is the need for assistance with their language. However, even within this population with a common need, there are major differences in their academic backgrounds. For example, ELL students differ in their level of English language proficiency (see, for example, Abedi et al., 2003). Some of these students are as proficient in English as their native English-speaking peers while some are transitioning from full proficiency in another language. For ELL students at the higher levels of English proficiency, accommodations such as “linguistically modified tests” might be more relevant whereas for those students at the lower level of English proficiency (but who have a higher level of native language proficiency), native language or bilingual assessments might be more relevant.

However, native language accommodations may not be relevant for all students who are proficient in their native language. Thus, a student’s academic background should be considered for assigning appropriate accommodations. For example, for bilingual students who are quite proficient in their native language but are being instructed in English, the native language or bilingual assessments may not be a relevant accommodation since they have learned all of the content terms in English (Abedi et al., 2004).

Another important background variable that could help with the appropriate assignment of accommodation is the experience students have with the accommodations in their instructional history. Students perform better under accommodations that they have used in their classrooms.

**Question 4: assessing the degree of relevance of accommodations for ELL students**

As indicated earlier in this chapter, the concept and application of accommodation was first introduced in the assessment literature for students with disabilities. There are many different
categories of disabilities with different background characteristics and different needs; therefore, accommodations should be relevant to each of these students in addressing their needs. For example, accommodations that are used for blind and visual impaired students (e.g., Braille) would not be relevant for deaf and hard of hearing students. For ELL students, however, the situation might be quite different. While their level of English proficiency might be different, all ELL students share a common need for assistance with their language. Therefore, many of the accommodations currently used for ELL students that are adopted from the field of students with disabilities are not relevant for these students. For example, accommodations such as “one-on-one testing,” “small group testing,” “multiple breaks,” “writing the response on the text booklet” may not have any relevance to ELL assessments. None of these accommodations specifically address their particular language needs. Providing such irrelevant accommodations may prevent students from receiving the accommodations that they actually need.

To assess the relevance of accommodations for ELL students in absence of credible experimental-based evidence, an accommodation expert panel at the state level could be formed to discuss the accommodations that are used at the state level and judge their relevance for ELLs. Willner et al. (2008) conducted an expert-group study to identify relevant accommodations for ELL students through a Delphi approach. The expert group selected a small set of accommodations that they recommend to use for ELL students.

**Recommendations for practice and future directions**

Accommodations are used for English language learners and students with disabilities to “level the playing field” by making assessments more accessible for these students. Unlike students with disabilities who are in different categories of disabilities with different needs, ELL students have a common need of language assistance. Therefore, accommodations that are created and used for students with disabilities for different purpose may not be relevant for ELL students. Willner et al. (2008) discussed direct-linguistic accommodations and indicated those accommodations may be more relevant for ELL students than the accommodations that are developed and used for students with disabilities. Unfortunately, there is not enough research support for many of the accommodations that are currently used for ELL students, even those that are under the category of “direct-linguistic accommodations.” The only way to judge the efficiency and validity of these accommodations is to use them in experimentally controlled situations on both ELL and non-ELL students and examine their validity, effectiveness, and differential impact under well-controlled experimental designs.

Based on the summary of research presented in this chapter, there is not enough evidence to help decide which accommodations to use for ELL students. In terms of research evidence, accommodations can be grouped into three categories. Category 1 includes accommodations that may alter the construct being measured. These accommodations are often referred to as “Modifications.” The outcome of assessments under these accommodations may not be aggregated with the non-accommodated assessment outcomes. As an example of these accommodations read-aloud test items in reading and providing full dictionary may be mentioned. For example, a dictionary may provide an unfair advantage to the recipient by helping them with the correct response to the questions. Category 2 includes accommodations for which research findings are mixed. For these accommodations more research is needed to shed light on their effectiveness and validity. Examples include extended time accommodations and providing glossaries. For accommodations in category 3, there is enough research support that could help to use them without much concern over their use. Linguistically modified tests are examples of these accommodations.
In summary, providing accommodations for ELL students helps not only to make assessments more accessible for these students but also to provide more valid assessments in general. However, in selecting accommodations for ELL students several major criteria must guide us. The purpose of accommodation is not to give unfair advantage to the recipients but to provide equal assessment opportunity to everyone; therefore, accommodations must be able to play this important role. In selecting accommodations one must clearly consider the purpose and use those accommodations that are relevant for such purpose.

Future developments

Issues concerning the use and validity of accommodations are gaining international momentum. Given the growth in the population of immigrants in the world and considering education issues of ELLs, accommodations are among the important concerns in the international education community. For example, it is likely that the United Kingdom may go down a route akin to the United States in coming years given the growing numbers of ELLs in state schools, and similarly in Germany with its Turkish or Kurdish communities. Educational researchers and policy makers have given their highest attention to these issues since accommodation issues are of major concern in the United States due to the increasing population of ELL students. There is much to be learned from the research amassed on accommodations in the United States as well as that from other countries facing the same situation.

Since accommodations are likely to become a major part of curriculum, instruction, and assessment in the future world’s educational system, examining their validity and effectiveness should have a high priority within each nation’s educational agenda. It is important to understand that ELLs have the knowledge and ability to learn the content knowledge at the same level as their native English-speaking peers, and accommodations should be used to “level the playing field,” i.e., give them opportunity to learn and opportunity to be fairly assessed as their peers as they are learning academic content in a language in which they may not still be quite proficient.

The international community can share and benefit from research on the accommodations for ELL students. It is therefore imperative to have an effective communication channel to collaborate with the research in accommodations and share the outcomes of such studies.

Acknowledgements

I wish to acknowledge the help of Kimberly Mundhenk and Nancy Ewers in reviewing and editing drafts of this chapter.

Further reading

Abedi, J. (2007). Utilizing accommodations in the assessment of English language learners. In N. H. Hornberger (ed.), Encyclopedia of Language and Education: Volume 7: Language Testing and Assessment. Heidelberg, Germany: Springer, 331–48. Concerns over the validity of accommodations for ELL students have surfaced as the international education community moves toward inclusion of all students (including ELLs and SDs) in national and local assessments. The results of accommodated and non-accommodated assessments cannot be aggregated if the validity of accommodations has not been established. This paper summarizes results of research on the effectiveness and validity of accommodations for ELLs and will examines research findings to determine if there is sufficient evidence to inform decision-makers on the provision of accommodations and reporting of accommodated assessments. While the main thrust of this paper is on accommodation issues for ELL students, some references have been made to accommodations for students with disabilities since ELL accommodation practices are highly affected by accommodation policies and practices for SDs.
Abedi, J., Hofstetter, C. and Lord, C. (2004) Assessment accommodations for English language learners: Implications for policy based research. *Review of Educational Research* 74: 1–28. Decisions about which accommodations to use, for whom, and under what conditions, are based on limited empirical evidence for their effectiveness and validity. Given the potential consequences of test results, it is important that policymakers and educators understand the empirical base underlying their use. This article reviews test accommodation strategies for English learners, derived from “scientifically based research.” The results caution against a one-size-fits-all approach. The more promising approaches include modified English and customized dictionaries, which can be used for all students, not just English language learners.

Francis, D., Rivera, M., Lesaux, N., Kieffer, M. and Rivera, H. (2006). Practical Guidelines for the Education of English Language Learners: Research-Based Recommendations for the Use of Accommodations in Large-Scale Assessments. (Under cooperative agreement grant S283B050034 for U.S. Department of Education). Portsmouth, NH: RMC Research Corporation, Center on Instruction. Available online at http://www.centeroninstruction.org/files/ELL3-Assessments.pdf. Language is a key component of individual student success yet English language learners are often unprepared for the rigors of academic language in content areas. This meta-analysis examines the effectiveness and validity of selected accommodations, familiar to the student through daily instructional use and applied during assessments. The results suggest the importance of linguistically appropriate accommodations, ones that have been used effectively in the classroom, and are appropriately selected to match individual student needs. Although no single accommodation has been shown to level the playing field, the most effective ones may be the use of a dictionary or glossary with extra time, provided that the student has previous classroom experience with dictionary or glossary use. Ensuring the opportunity to learn during instruction combined with accommodations that are familiar and useful to individual students in the classroom can increase the chance of academic success for all students.

Kieffer, M., Lesaux, N., Rivera, M. and Francis, D. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research* 79: 1168–1201. Accommodation use is an important consideration in improving assessments for English language learners and warrants continued research. However, these researchers hypothesize that the performance gap between English language learners and native speakers of English may be due to the impact of necessary language skills rather than unnecessary language demands which accommodations address. They suggest that academic language and content knowledge intersect and are highly dependent on classroom opportunities to learn, without which accommodation use will not result in improved academic performance. They recommend that assessment accommodation research continue but that a strong focus is placed on instruction and academic language that is so important for academic success.

Sireci, S., Li, S. and Scarpati, S. (2003). *The Effects of Test Accommodation on Test Performance: a review of the literature* (Report No. 485). Amherst, MA: School of Education, University of Massachusetts, Center for Educational Assessment. These authors call into question the interaction hypothesis that only students who need an accommodation benefit from it, and those for whom it is not intended will not. The focus on four accommodations: customized dictionaries or glossaries, extra time, test translation/adaptation into native language translations, and simplified English accommodations results in mixed findings. The authors attribute the inconclusive results to the fact that both accommodations and groups of students are heterogeneous. They further discuss the four accommodations in detail together with potential in the concept of Universal Design for Learning (UDL) to increase test accessibility for the widest number of students; possibly reducing or eliminating the need for accommodations. Research in the challenging arena of accommodations has produced positive results for many groups of students and will continue to improve assessment practices.

**References**


