The Use of Pictorial Supports as an Accommodation for Increasing Access to Test Items for Students with Limited Proficiency in the Language of Testing

Guillermo Solano-Flores
University of Colorado at Boulder

Paper presented at the 7th Conference of the International Test Commission.
Hong Kong, July 19-21, 2010.

Abstract

This paper reports on an NSF-funded project that examines vignette illustrations (VIs) as a form of testing accommodation for English language learners (ELLs)—students who are developing English as a second language yet they are tested in English, in major assessment programs in the U.S. VIs are pictorial supports intended to make the content of test items more accessible to ELLs without altering their text and without giving away their answers. We have developed a procedure for systematically designing VIs. Based on semiotics, socio-cultural theory, and cognitive science, our procedure allows identification of both linguistic/cultural challenges—constituents (words, phrases, terms, idiomatic expressions) which may pose challenges to ELLs due to their limited English proficiency or their limited experience with certain contextual information)—and linguistic/cultural affordances (constituents that are not likely to pose these challenges to ELLs). Based on the identified linguistic and cultural challenges and affordances, illustration development teams composed by bilingual teachers, science teachers, and science content experts, write scripts that specify the characteristics that the illustrations should have. The paper discusses the procedure for developing VIs and discusses the potential of VIs as a valid, cost-effective, easy-to-implement testing accommodation in multilingual and multicultural contexts in which student language proficiency in the language of testing is a potential threat to test validity.

In the United States, legislation mandates the inclusion of English language learners (ELLs) in large-scale testing programs after a short time of being schooled in English. This practice is inconsistent with a large body of evidence from the field of language development and second language acquisition, which shows that, although ELLs can develop basic communication skills in their second language in a short time, developing their academic language in that second language takes considerably more time (Hakuta & Beatty, 200; Hakuta, Butler, & Witt, 2000; Katz, Low, Stack, & Tsang, 2004).

This practice is partially based on the unrealistic expectation about the effectiveness of testing accommodations. Testing accommodations for ELLs are modifications made on the ways in which tests are administered with the intent to make the content of the items more accessible to these students without giving away the correct answers of test items and without giving them an unfair advantage over the non-ELL students tested without accommodations (Abedi, Lord, Hofstetter, & Baker, 2001). There are multiple forms of accommodations authorized by major national and state assessment systems used in the U.S. (Rivera, Collum, Willner, & Sia, 2006). They include, among many others, providing ELLs with translations of the tests, allowing them to answer test items in their own language, reading the items aloud for them, and simplifying the

wording of items. Most of these accommodations target one of the four language modes--listening, speaking, reading, or writing.

Unfortunately, many of these accommodations lack empirical or theoretical support that justifies their use. Also, their effectiveness may limited by the fact that, since each ELL has a unique schooling history in English and a unique set of strengths and weaknesses in listening, speaking, reading, and writing in English, only few students can benefit from a given form of accommodation (see Abedi, Lord, Hofstetter, & Baker, 2001). Finally, while some forms of accommodation are potentially effective, they are difficult or unlikely to be properly implemented (Solano-Flores & Trumbull, 2008).

Several years ago, while examining the responses of ELL students to items from the NAEP (National Assessment of National Progress—a national assessment administered in the U.S.; see Solano-Flores, Li, Speroni, Rodriguez, Basterra, & Dovholuk, 2007), we observed that higher percentages of ELL students responded correctly to mathematics items when they were accompanied by illustrations.

The finding was especially intriguing because, in some cases, the illustrations did not appear to be relevant to understanding the items. For some of them, the illustration was not referred to by the text of the item (e.g., it did not ask the student to look at the illustration nor the illustration provided any information that was indispensable to understanding the item or responding to it).

In looking for empirical evidence on the use of illustrations in testing, we found that, while illustrations are a frequent form of device used in science textbooks and science tests, with very few exceptions, illustrated items are a neglected topic in research in testing. In addition, no research on illustrated items has been conducted to examine its possibilities in the testing of ELLs. Moreover, state and national assessment programs in the U.S. do not make any provision for the use of illustrations as a form of testing accommodation for ELLs.

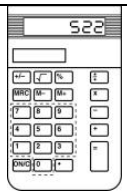
The Concept of Vignette Illustration

Two years ago, we obtained funding from the National Science Foundation to investigate these matters more carefully. In our investigation, we focus on a specific form of item, multiple-choice item. Also, we focus on a specific form of illustration that we call, *vignette illustration*—a pictorial support added to an item without altering its text and with the intent to making its content more accessible to ELL students but without giving away its response. Thus, a vignette-illustrated item can be defined as an illustrated item with three properties (see Figure 1):

1) The illustration provides a simple, concrete representation of one of the constituents (i.e., a word, term, phrase, or expression).
2) The text of the item does not refer the test taker to the illustration; it does not direct the student to examining the illustration nor the illustration provides any information that is indispensable to understanding the item or responding to it.
3) The text of the item provides all the information needed to understand it and respond to it; removing the illustration does not affect the coherence of the information provided by the text of the item.

Because vignette illustrations do not alter the text of items and their appropriate use does not depend on the skills of those who administer tests, they have the potential of being an efficient and cost-effective form of testing accommodation for ELLs.

Item 1



Mrs. Jones bought 6 pints of berries. Each pint cost 87¢. Mrs. Jones used her calculator to find the cost of the berries and the display showed 522. What was the cost of the berries?

A) $522
B) $52.20
C) $5.22
D) $0.52

Figure 1. An example of vignette-illustrated item. Source: National Assessment of Educational Progress. (1996). *Mathematics items public release*. Washington, DC: Author.

Goals

Our investigation addresses three research questions:

1. *What principles underlie the effective design of science and mathematics test items with illustrations in ways that minimize limited English proficiency as a factor that prevents ELLs from understanding the items?*
2. *Is the presence of an illustration a moderator in students' understanding of test items?*
3. *Does the presence of an illustration have a different effect on the performance of ELLs and the performance of non-ELL students?*

This paper focuses on the activities conducted to address the first research question. We aim to identify a set of principles for the proper design and use of vignette illustrations as a form of testing accommodation for ELLs. To meet that purpose, we have developed a methodology for developing and using vignette-illustrated items.

Theoretical Framework

Semiotics, cognitive psychology, linguistics, and socio-cultural theory are brought together to examine the complexity of image and text interpretation. Especially important to our work has been the theoretical work and empirical evidence provided by Richard Mayer and his associates (e.g., Mayer & Sims, 1994). Evidence from those fields indicates that text and visual images can be seen as two independent but interacting sources of information which are processed independently and integrated by the reader-viewer. If the information from the two sources is conflictive, or if the information presented visually is too complex, the illustration will not meet the purpose with which it was created. In other words, illustrations should not be assumed to be effective in providing information unless they parallel text without being distractive (Filippatou and Pumfrey, 1996; Harp & Mayer, 1997, 1998; Mayer, Heiser & Lonn, 2001).

Developing Vignette Illustrations for English Language Learners

During the last two years, my colleagues and I have developed and perfected a procedure for developing vignette illustrations. The procedure is shown in Figure 1. Five aspects are discussed here: 1) the team of developers that need to participate in the process of development of vignette illustrations, 2) the analysis of linguistic and cultural challenges and affordances that an item is likely to pose to ELLs, 3) scripting, 4) the setting of illustration parameters, and 5) the notion of illustrability.

Illustration development team. Three types of professionals need to work in developing vignette illustrations: bilingual teachers, science teachers, and content experts (scientists). Bilingual teachers provide first-hand knowledge on the language usage and the culture of the target population. Science teachers provide knowledge of the topics assessed and the level of complexity of images that is appropriate to the grade level of the population tested. Content experts examine the accuracy of the illustrations.

Linguistic and cultural challenges and affordances. For each item, with facilitation form project staff, the illustration development team identifies the constituents that are likely to pose a linguistic challenge to ELLs or have cultural connotations that are likely to be unfamiliar to these students. It also identifies those constituents that are familiar to students and therefore, are unlikely to be challenging to ELL students. Also, it identifies the constituents that are key to the construct being assessed and which, consequently, should not be illustrated.

Scripting. From the analysis of linguistic and cultural challenges and affordances, the illustration development team writes a script intended to be used by illustrators to create the vignette-illustrations. The script specifies the characteristics of the illustration. This script is based on the constituents from the stem, not the options of the item. However, as part of developing the script, the item development team needs to make sure that it does not give away the correct answer of the multiple choice item or mislead the student in selecting a distractor. Along with the script, the illustration development team provides illustrators with reference materials that help them to understand the specifications.

Illustration parameters. Critical to effective scripting is the use of illustration parameters that define the characteristics that the illustrations to be developed should and should not have. These illustration parameters are set up by the project staff according to the characteristics of the population of examinees. In our project, they have been established based on existing knowledge on the cognitive complexity of text and image interpretation (see conceptual framework above). According to this knowledge, item vignette illustrations should not be expected to have the same set of characteristics of other forms of illustrations, such as those typically used in science textbooks. For example, whereas science textbooks frequently include illustrations that show sequence of events or actions (procedures) or events (processes), those functions cannot be attained in illustrations that are intended to supplement the information of the item at a glance, in a non-distractive manner. Table 1 presents the parameters the we have established to guide the scripting of illustrations for vignette-illustrated items for ELLs.

Table 1. Parameters used in the design of vignette illustrations for English language learners.

Illustration features allowed in the project

- a specific object or event mentioned in the stem
- an example of an idea or set of cases
- features or attributes common to a given class of objects
- a comparison or contrast of objects
- a situation, event, or set of circumstances
- the parts that comprise a whole
- select, basic components of something

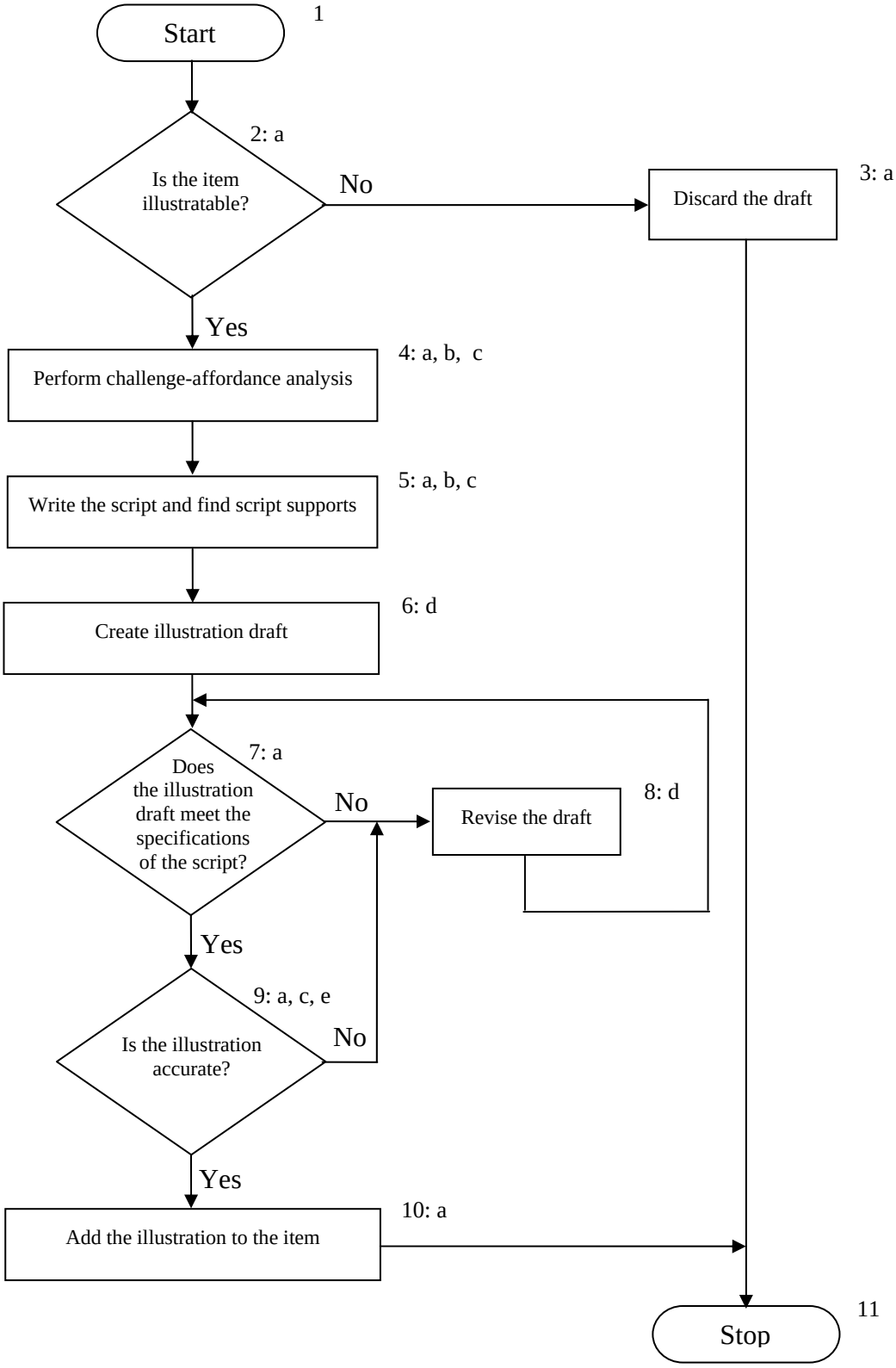
Illustration features not allowed in the project

- an object or action mentioned in an option
- multiple stages
- multiple actions
- objects in different scales
- sequence of events, stages of a process or procedure
- taxonomical or hierarchical relationships between objects
- references to experiences accessible only to some individuals (e.g. private jokes)
- images intended to produce certain attitudes or moods
- symbolic representation
- inserts
- arrows
- zoom in or zoom out beyond student's immediate personal everyday life experience (naked eye)

Graphic design specifications

- Color: Black and white.
- Tone: No gray tones
- Size: 2.5 inch x 2.5 inch
- Position of the illustration: To the right of the item.
- Framing: No framing.
- Inserts. No inserts.
- Realism: Realistic.
- Text: No labels.
- Symbols: No symbols.

Figure 2. Procedure for developing test item vignette illustrations for ELLs.



Illustrability. Not all the items in a test can be illustrated. We estimate that about 40% of the items in tests used in large-scale assessment programs do not lend themselves for illustration (Wang & Solano-Flores, 2010). We have identified that an item is unillustratable because the item does not provide contextual information, does not refer to any concrete object or event ("photosynthesis"), refers to objects or events that cannot be seen with naked eye, or refers to objects or events that are not ("Which of this is not..."). An item is also unillustratable if an illustration of it leads students to selecting any of the distractors, gives away the correct answer, uses or includes multiple objects or sequence of objects, or needs to show objects or events that are not.

Summary and Final Remarks

Several reasons make the use of vignette-illustrated items in the testing of ELLs worth investigating. This form of accommodation does not require modifying the wording of items—which reduces the possibility of altering the construct being measured. Also, while this form of accommodation may require the use of certain principles for its proper design, the fidelity of its implementation does not depend on the skills of test administrators. Moreover, the use of vignette illustrations is not based on any assumptions on the degree of proficiency of students in English or in their native languages. Furthermore, this form of testing accommodation appears to be cost-effective.

We intend to produce in the near future a set of documents and procedures for the systematic and cost-effective design and analysis of vignette-illustrated items (see Solano-Flores & Wang, 2010). These documents will target test developers and educators, and will inform the process of test development in both the context of large-scale testing and the context of classroom assessment.

We are currently developing the vignette illustrations for test items from the Colorado Student Assessment Program (CSAP) and Trend in Mathematics and Science Study (TIMSS). To evaluate the effectiveness of vignette illustrations as a form of accommodation for ELLs, we will administer ELLs and non-ELLs illustrated and non-illustrated versions of the same set of science items from those tests. A series of analyses of variance will allow us to determine any statistically significant score differences in favor of the illustrated version of the items. The analyses also will allow us to assess the extent to which the accommodation contributes to reduce the amount of measurement error due to language and to reduce the score gap between non-ELL and ELL students.

As an additional analysis, the transcriptions of cognitive interviews conducted with another sample of students from the same population of ELLs, and who were given illustrated and non-illustrated versions of the items are currently being coded and analyzed. Results from this qualitative analyses will produce empirical data on the cognitive validity of vignette illustrations.

The effectiveness of vignette illustrations in the testing of ELLs is yet to be determined based on the empirical evidence that will be obtained with the research design described. In the meantime, we believe that we have successfully completed the first (and probably the most difficult) part of the investigation—creating a procedure for systematically developing this form of accommodation.

Author's Note

The investigation described is funded by the National Science Foundation, Award DRL 0822362. I wish to thank my colleagues Lucinda Soltero-Gonzalez, Chao Wang, and Rachel Prosser for their participation and to our team of bilingual and science teachers for their

enthusiastic support and hard work. The opinions here expressed are not necessarily those of my colleagues or the funding agency.

References

Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2001). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice, 19(3)*, 16–26.

Filippatou, D. & Pumfrey, P.D. (1996). Pictures, titles, reading accuracy and reading comprehension: A research review (1973-95). *Education research*, *38*(3), 259-291.

Hakuta, K. & Beatty, A. (2000). (Eds.) *Testing English-language learners in US Schools.* Washington, D.C. National Academy Press.

Hakuta, K., Butler, Y.G., & Witt, D. (2000). *How long does it take for English learners to attain proficiency?* Policy Report 2000-1. Santa Barbara, CA: University of California Linguistic Minority Research Institute.Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89*, 92-102.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*, 414-434.

Katz, A., Low, P., Stack, J., & Tsang, S.-L. (2004). *A study of content area assessment for English language learners*. Final Report prepared for Office of English Language Acquisition and Academic Achievement for Limited English Proficient Students. Oakland, CA: ARC Associates.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extension of a dual-coding theory of multimedia learning. Journal of Educational Psychology, 86(3), 389-401.

Rivera, C., Collum, E., Willner, L. N. & Sia, J. K. (2006). Study 1: An analysis of state assessment policies regarding the accommodation of English language learners. In Rivera, C. & Collum, E. (Eds.), *State assessment policy and practice for English language learners: A national perspective*. Mahwah, NJ: Lawrence Earlbaum Associates, Publishers. (pp. 1-136).

Solano-Flores, G., Li, M., Speroni, C., Rodriguez, J., Basterra, M., & Dovholuk, G. (2007). *Comparing the properties of teacher-adapted and linguistically-simplified test items for English language learners*. Paper presented at the annual meeting of the American Educational Research Association. Chicago, IL. April 9-13.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher*, *32*(2), 3-13.

Solano-Flores, G., & Wang, C. (2010). *Development and use of a conceptual framework for analyzing and classifying illustrations used in science assessment*. Paper proposal submitted to the Annual Conference of the National Council on Measurement in Education, New Orleans, LA, April 7-11, 2011.

Wang, C., & Solano-Flores, G. (2010). *The use of Illustrations in large-scale science assessment: A comparative study*. Paper proposal submitted to the Annual Conference of the American Educational Research Association, New Orleans, LA, April, 2011.