Potential Benefits of Bilingual Constructed Response Science Assessments for Understanding Bilingual Learners' Emergent Use of Language of Scientific Investigation Practices

Cory Buxton, Martha Allexsaht-Snider, Rouhollah Aghasaleh, Shakhnoza Kayumova, Seo-hyun Kim, Youn-Jeng Choi, and Allan Cohen
University of Georgia

Background and Rationale

School-determined English as a Second Language (ESOL) levels may not be the best measure for researchers or teachers who wish to understand students' functional usage and comfort with their home language and the language of instruction in classroom contexts (Solano-Flores, 2008). There are a number of psychological and sociological factors, as well as questions of prior academic background in and support for the home language, that influence the linguistic choices that students make regarding academic communication (Levine, 2011). By providing a bilingual constructed response assessment and tracking the language(s) in which students both read the questions and wrote their responses (combined with basic demographic information), we attempted to create a more useful functional measure of student ability in using home and school languages for academic purposes (Shaw, Bunch, & Geaney, 2010). Such a resource could be used to support teachers in exploring the potential benefits of building on first-language resources to support science learning through writing (Rosebery & Warren, 2008).

We conceptualize science learning as mastering a combination of practices, core conceptual ideas, and communication skills that are developed across a broad range of life-wide learning contexts (National Research Council, 2011, 2009). From this perspective, a primary goal of science learning is to gain competence in applying critical thinking processes used for problem solving during science tasks (Kuhn, 2005). For Kuhn, these critical thinking processes include generating and testing hypotheses, using cause-and -effect reasoning, learning to control variables, and arguing from evidence. We extend Kuhn's ideas by focusing on science tasks that prompt students' science talk, writing, and action in ways that support development of critical communication practices as well as critical thinking practices. Duschl and Grandy (2008) have argued that worthwhile science learning tasks make student thinking visible to help teachers improve students' scientific thinking. Writing about science investigations can serve this purpose (Gunel, Hand, & Prain, 2007).

We developed a bilingual constructed response assessment that asks students to write about science investigations (Buxton et al., 2013) as part of our Language-Rich Inquiry Science with English Language Learners (LISELL) Project to support a more robust measure of students' scientific thinking as expressed through writing. There is a limited but growing body of research on the value of teachers attending to the science writing of emergent bilingual learners as a way to both understand and improve these students' science learning experiences. For example, Huerta, Lara-Alecio, Tong, and Irby (2014) developed and tested a science notebook rubric to measure both the conceptual understanding and the academic language skills expressed in upper elementary grade second language learners' writing. Similarly, Furtak and Ruiz-Primo (2008) compared middle school students' written responses with their oral discussions about the same formative assessment prompts and found that the written responses elicited a greater range of students' science conceptions. Further, one of us was involved in an earlier study that used emergent

DOI: 10.37514/DBH-J.2014.2.1.05

bilingual learners' writing to measure their developing English proficiency and ability to explain science concepts by exploring the relationship between the form (i.e., conventions, organization, and style/voice) and content (i.e., specific knowledge and understanding of science) of their expository science writing (Lee, Penfield, & Buxton, 2011). Thus, while there is growing evidence on the value of studying bilingual learners' writing for supporting their science learning, there is less evidence about how best to support these students' science writing on either formative or summative assessments as part of that process.

The research to date on language accommodations for bilingual learners on assessments is both limited and ambiguous, with a wide range of potential accommodation strategies having been proposed to support emergent bilingual learners. For example, during the 2005 and 2009 National Assessment of Educational Progress (NAEP) science assessments in the United States, English learners were allowed the use of bilingual dictionaries, bilingual texts, subject-specific glossaries, English language dictionaries, extra time to complete assessments, having test items read aloud, permission to ask clarifying questions, and permission to dictate oral responses to a scribe (National Center for Education Statistics [NCES], 2011). All of these accommodations were meant to help English learners make sense of the language of the assessment as written. None of the accommodations, however, modified the language of the assessment itself, such as through translation of test items, or simplification of language to avoid unnecessarily complex grammatical constructions, polysemic terms, or idiomatic expressions.

In a meta-analysis of testing accommodations, Kieffer, Lesaux, Rivera, and Francis (2009) examined the effectiveness and the validity of specific testing accommodations for improving the performance of bilingual learners on large-scale assessments in mathematics and science. Effectiveness was defined in terms of whether bilingual learners who received an accommodation outperformed bilingual learners who did not receive the accommodation. Validity was defined in terms of whether an accommodation altered the overall construct validity of a test, measured by whether monolingual learners who received an accommodation outperformed monolingual learners who did not receive the accommodation. The researchers analyzed data from 11 studies on the effects of accommodations for bilingual learners on NAEP, Trends in International Mathematics and Science Study (TIMSS), and various state assessments in the United States in mathematics and science at fourth and eighth grades. The studies considered accommodations including extra time, dual language questions, a Spanish version of the assessment, use of a bilingual dictionary, use of an English dictionary, and simplified English language versions of test questions. The results indicated that the only accommodation that consistently improved bilingual learners' performance was providing students with an English language dictionary. This finding raises intriguing questions about why various first language supports and other linguistic accommodations did not yield improved test performance.

In a contrasting finding, Siegel (2007) studied accommodations for bilingual learners on teacher-developed assessments in two seventh-grade life science classes that used National Science Foundation supported science curriculum materials. Participating students were bilingual learners who had some proficiency in reading and writing in English, as well as native English speakers in the same classes. Siegel studied 11 kinds of assessment modifications made to support linguistic (e.g., reducing the number of words), cognitive (e.g., adding a graphic organizer), and visual (e.g., including a picture) clarifications to assessment items. Results indicated that both native English speakers and emergent bilingual learners scored significantly higher on the modified classroom assessments. Thus, these modifications, unlike the accommodations studied in Keiffer et al's analysis of standardized tests, did improve bilingual learners' performance but also caused validity threats, in that the performance of native English speakers increased as well.

Part of the ambiguous nature of these findings likely has to do with students' comfort and experience using their home language in the context of schooling. At least two factors may contribute to this comfort. First, there is an academic component, including whether students possess sufficient general academic and technical science vocabulary and broader experience reading and writing the language of science in their home language (Duff, 2011). Students may be more conversationally fluent in their home language but may not possess sufficient academic language skills in that language to feel comfortable using it in an academic context if they have received little or no home language support in learning the academic material (Sotelo-Dynega, Ortiz, Flanagan, & Chaplin, 2013). Second, there is a social component, including the broader question of whether students feel that it is acceptable to use their home language in the school context for academic purposes (Hoff, 2006). If students have received the message, either explicitly or tacitly, that their home language does not belong in school, or is only appropriate for social interactions, then they are likely to feel uncomfortable using this language during an academic task, such as an assessment (Wong Fillmore, 2000). When taken together, these academic and social factors may all but cancel out any potential benefit of a bilingual assessment in helping teachers to get an accurate picture of what students understand about a topic. Thus, our inclusion of bilingual science activities and bilingual academic discussions with teachers, students, and families as part of the broader LISELL professional learning experience was meant to address both of these challenges by increasing exposure to and comfort with academic language in Spanish and by making it clear to students and teachers that a student's home language is a valuable resource that should be fostered and welcomed in the science classroom (Buxton, Allexsaht-Snider, & Rivera, 2012). We hoped that given these messages, students would be more likely to make use of the resources of a bilingual assessment.

Conceptual Framework

Systemic Functional Linguistics (SFL) is a theory of language that argues that our experience of things is driven by analogies and metaphors, or in other words, how things are like other things (Eggins, 2004). From this perspective, a primary role of language, and of grammar in particular, is to clarify how new things we experience fit in with our prior understandings (Halliday, 1994). Thus, as young children learn language, that language both shapes and is shaped by their growing experience with the world around them.

As children move toward school age, and their language skills continue to develop, they begin to interact with language in new ways, through reading and writing. As Halliday (2004) noted,

> We think of writing as just…a new medium: children already know the fundamentals of language; they're now going to learn to process language visually in order to gain access to books and magazines and all the other trappings of our written culture. We don't think of [this]…as changing the way they mean. (p. 12)

But in fact, learning to read and write does fundamentally change the way we experience the world and the way we make sense of those experiences, and not just by broadening the experiences to which we gain access. The grammar of written text also changes our experience of the world. As SFL theory has illuminated, written text, at least in Western languages, often transforms the names of happenings (verbs) into things (nouns), such that *create* becomes *creation* and *measure* becomes *measurement*. Halliday (1994) argued that this trend began with ancient Greek philosophy and the philosophers' desire to create abstract objects that were more persistent, and therefore easier to categorize, than were transient happenings. Thus, from the SFL perspective, the value and the power of grammar come from its ability to turn one linguistic class of words into another, what Halliday refers to as grammatical metaphor.

The grammatical transformation of linguistic classes that began with philosophy became even more important for the purposes of science during the Enlightenment. Galileo, Newton and others embraced the power of grammatical metaphor to create our modern discipline-based organization of technical knowledge, and along with it, our modern language of science (Chomsky, 2004). This language of science, like the language of philosophy, requires more stable and persistent linguistic classes for the purpose of creating technical taxonomies. To this end, the language of science employs a grammatical drift from circumstances (unstable and ephemeral) through processes and qualities, toward entities that are stable and persistent through time (Halliday, 1999). While we may not routinely think about the language we hear in terms of its grammatical function, we can easily distinguish language we encounter as sounding more or less "scientific" to the ear. Below, we exemplify this process, using samples from our students' writing on one of our assessment items dealing with a pond ecosystem:

1. "The big fish have nothing to eat and die." This is an example of a *circumstance*–it is grammatically unstable and temporary, and we do not perceive the language as scientific.

2. "If the population of little fish decreases, then the population of big fish decreases too." This example demonstrates a *process*–it is somewhat more stable in time, and therefore sounds more scientific.

3. "Diseased small fish cause a decrease in the number of big fish." This example, while expressing the same idea as the prior example, grammatically shifts to demonstrate a *quality* of the small fish (they are *diseased*) that is more persistent and thus sounds still more scientific.

4. "Population density of the large fish depends on the density of the small fish population." In this case, a stable *entity* has been grammatically constructed (*population density*) and typifies the fully formed language of science.

As Halliday (2004) has explained it, the functional goal of the language of science is to express how "experimental science theorizes the physical world of processes by creating a new world populated largely by virtual entities created at the intersection of process and thing" (p. 44).

How do these features of the language of science influence students' experience of the world around them? Upon entering middle school, many students are systematically exposed to the grammatical structures of the science disciplines for the first time. Students are suddenly asked to contextualize and interpret their experiences of the natural world through the language of science, which may sound quite foreign to many of them. This technical language is critical for success in academic science because it allows for the development of sustained scientific discourse needed to engage in the science and engineering practices that are a centerpiece of the Next Generation Science Standards in the United States (Achieve, 2013).

While the technical nature of the language of science, as explained through functional linguistics, is a new and challenging discourse for many students, it is doubly challenging for emergent bilingual learners who must confront this discourse in a second (or even a third) language. Classroom materials that mistakenly take students' familiarity with academic discourse structures for granted are clearly assessing more than students' understanding of science concepts. The shift from everyday to scientific understanding of the natural world can be seen as parallel to the shift from spoken to written language. Grammatical drift of the kind discussed above makes language, even when spoken, sound more like its written counterpart. Thus, practice writing the language of science will enhance the same skills needed to speak the language of science. To this end, our development of a bilingual constructed response science assessment as part of the LISELL project was designed to help us better understand how emergent bilingual learners make linguistic choices and how to guide teachers in helping their students to better express their scientific thinking through the language of science, both orally and in writing.

Methods

Data Collection

In this study, we used data sampled from student assessments collected during the 2011-2012 academic year. The assessments were given as part of the broader LISELL research and development project. The overarching goals of this project were to develop, implement, and revise both a pedagogical model and a professional learning framework for improving science teaching and learning for all middle school students, with a special focus on emergent bilingual learners. The LISELL pedagogical model was developed to promote students' science talk, writing, and action in ways that support development of practices central to scientific investigation, what we refer to as the *language of scientific investigation practices*. Specifically, the LISELL pedagogical model encompasses five practices that are meant to help all students and particularly emergent bilingual learners to develop proficiency using the language of science: (a) coordinating hypothesis, observation and evidence, (b) controlling variables to design a fair test, (c) explaining cause and effect relationships, (d) developing general academic vocabulary in context, and (e) owning the language of science.

As part of our research on this pedagogical model, we developed our bilingual constructed response assessment, in which students were asked to write responses to a series of scientific investigation scenarios. The topics of these scenarios, such as lifting weights, growing plants, rolling balls, and observing shadows, were selected to incorporate both students' lived experiences and core science concepts. The goal of the assessment was to measure how well students could apply the science practices of the LISELL pedagogical model in the context of science experiments using the language of science. We built into our assessment a number of the linguistic, cognitive, and visual clarifications that Siegel (2007)–discussed above–found to be helpful in improving the performance of emergent bilingual learners on classroom assessments.

In addition to being bilingual in English and Spanish, our assessment items were attentive to word choice to reduce unnecessary academic language, included relevant diagrams and pictures for each question, and used contexts and scenarios for the questions that middle school students should have some experience with outside of school. There were two versions of the assessment–a pink form and a blue form–each of which consisted of six constructed response questions: two questions each on the topics of coordinating hypothesis, observation, and evidence; controlling variables to design a fair test; and explaining cause and effect relationships. Questions on each form were meant to assess the same constructs but using different scenarios. The assessments were given to all students in the classes of the 16 science teachers who participated in our professional learning activities during 2011-2012. Half of the students were randomly given the blue form and the other half of students the pink form during the pre-assessment. For the post-assessment, each student was given the opposite form from the one taken during pre-assessment.

During the 2011-2012 school year, a total of 1,714 students took our project assessment at the start of the year and 1,552 students took the assessment at the end of the year, with 1,270 students taking both the pre- and post-assessments. On the post-assessment, we asked students to indicate if they self-identified as Latino/a or not. A total of 666 students (42.9%) identified as Latino/a, of whom 546 students had also completed the pre-assessment. Thus, we had a data set of 546 self-identified Latino/a students who had completed the LISELL pre- and post-assessments in 2011-2012. We note that nearly all Latino/a students in the settings where we work are first or second generation newcomers who speak Spanish as their home language, while all academic instruction in our project schools was in English only. These students' proficiency in English varies widely, as does their proficiency in academic Spanish. We asked students to indicate on the assessment if they read each question in English only, in Spanish only, or in both English and Spanish. Additionally, we looked to see if students' written responses on the assessment were in English only, in Spanish only, or in both languages. On the pre-assessment, 41 of the 546 Latino/a students (7.5%) wrote in Spanish for at least part of the assessment, 208 of the 546 Latino/a students (38.1%) read at least some of the questions in both Spanish and English, but wrote their responses in English only, and 297 students (54.4%) read and wrote in English only on all questions. On the post-assessment, 8.8% of the self-identified Latino/a students wrote in Spanish for at least part of the assessment, 41.7% read at least some questions in both Spanish and English, but wrote their responses in English only, and 50.5% read and wrote in English only on all questions.

For this study, we limited our analysis to the assessments of Latino/a students, because we were interested in the language choices of Spanish-speaking bilingual students. We analyzed both pre- and post-assessments, but did not match pre- and post-assessments for particular students. Thus, we did not attempt to assess individual student growth, but rather to identify general language use patterns for Latino/a students across the range of classrooms and schools in our study. We did not consider gender in our selection criteria, but because we selected randomly from a pool that was roughly equal in gender distribution, our samples likewise came out to be roughly gender balanced (see Table 1). Likewise, we did not consider grade level (6^{th}, 7^{th} or 8^{th}) or school (students were from five different schools) when selecting student assessments. These distributions are shown in Table 1 for each of our language use criteria groups.

We randomly selected 30 Latino/a students' assessments (15 of the blue form and 15 of the pink form) from the total pool of assessments for each of the following six language use criteria:

- student wrote in Spanish only for at least one question on pre-assessment;
- student read in both languages for at least one question but wrote in English only on pre-assessment;
- student read and wrote in English only on pre-assessment;
- student wrote in Spanish only for at least one question on post-assessment;
- student read in both languages for at least one question but wrote in English only on post-assessment;

- student read and wrote in English only on post-assessment.

This yielded a total of 180 student assessments. We analyzed three of the six items on each form of the assessment (blue and pink) to include each of the three LISELL science investigation practices: coordinating hypothesis, observation, and evidence; controlling variables to design a fair test; and explaining cause and effect relationships (sample questions are included in the appendix). Thus, on each of the 180 randomly selected student assessments, we analyzed three different questions, for a total of 540 student responses analyzed.

Table 1 *Distribution of Assessments by Gender, Grade Level and School*

Language Group	Gender	Grade Level	School
Written in Spanish, pretest Blue	Male = 6 Female = 9	6^{th} = 5; 7^{th} = 5; 8^{th} = 5	A= 2; B= 6; C= 1; D= 2; E= 4
Written in Spanish, pretest Pink	Male = 7 Female = 8	6^{th} = 5; 7^{th} = 6; 8^{th} = 4	A= 3; B= 4; C= 5; D= 1; E= 2
Written in Spanish, posttest Blue	Male = 6 Female = 9	6^{th} = 7; 7^{th} = 3; 8^{th} = 5	A= 2; B= 4; C= 3; D= 2; E= 4
Written in Spanish, posttest Pink	Male = 5 Female = 10	6^{th} = 7; 7^{th} = 5; 8^{th} = 3	A= 1; B= 4; C= 1; D= 6; E= 3
Written in Spanish, Totals	***Male = 24 Female = 36***	***6^{th} = 24; 7^{th} = 19; 8^{th} = 17***	***A= 8; B= 18; C= 10; D= 11; E= 13***
Read in Both, pretest Blue	Male = 7 Female = 8	6^{th} = 5; 7^{th} = 7; 8^{th} = 3	A= 2; B= 4; C= 2; D= 3; E= 4
Read in Both, pretest Pink	Male = 8 Female = 7	6^{th} = 4; 7^{th} = 4; 8^{th} = 7	A= 2; B= 5; C= 2; D= 5; E= 0
Read in Both, posttest Blue	Male = 9 Female = 6	6^{th} = 4; 7^{th} = 7; 8^{th} = 4	A= 5; B= 4; C= 2; D= 4; E= 0
Read in Both, posttest Pink	Male = 7 Female = 8	6^{th} = 5; 7^{th} = 4; 8^{th} = 6	A= 1; B= 4; C= 5; D= 3; E= 2
Read in Both, Totals	***Male = 31 Female = 29***	***6^{th} = 18; 7^{th} = 22; 8^{th} = 20***	***A= 11; B= 17; C= 11; D= 15; E= 6***
Read in English only, pretest Blue	Male = 11 Female = 4	6^{th} = 2; 7^{th} = 9; 8^{th} = 4	A= 2; B= 3; C= 2; D= 6; E= 3
Read in English only, pretest Pink	Male = 6 Female = 9	6^{th} = 5; 7^{th} = 6; 8^{th} = 4	A= 2; B= 2; C= 4; D= 5; E= 2
Read in English only, posttest Blue	Male = 7 Female = 8	6^{th} = 5; 7^{th} = 6; 8^{th} = 4	A= 5; B= 2; C= 3; D= 3; E= 2
Read in English only, posttest Pink	Male = 8 Female = 7	6^{th} = 3; 7^{th} = 8; 8^{th} = 4	A= 2; B= 0; C= 8; D= 3 E= 2
Read in English Only, Totals	***Male = 32 Female = 28***	***6^{th} = 15; 7^{th} = 29; 8^{th} = 16***	***A= 10; B= 7; C= 17; D= 17; E= 9***
Grand Totals	**Male = 87 Female = 93**	**6^{th} = 56; 7^{th} = 71; 8^{th}= 53**	**A= 29; B= 42; C= 38; D= 43; E= 28**

A few notable demographic patterns can be seen in Table 1. First, girls in our sample were more likely to write in Spanish than were boys (by a 3:2 ratio). The gender breakdowns were approximately even for the other two language categories (reading in both languages and in English only). Second, the youngest students (sixth graders) were more likely to write in Spanish, while as students grew older, the number of students writing in Spanish decreased relative to their proportion in the sample. Finally, two schools, School B and School E, had a notably higher percentage of their students write in Spanish, while two other schools, School C and School D, had notably higher percentages of their students write in English only.

Data Analysis
In conducting this functional language analysis, we consider two conceptual perspectives on the language of science. First, from the systemic functional linguistics (SFL) perspective, we consider the language of science as having unique qualities in terms of how the grammar of student responses is constructed. Following both Halliday (2004) and Fang and Schleppegrell (2008), we analyzed students' responses in terms of (a) technical vocabulary usage, (b) drift toward more stable linguistic classes, (c) lexical density, and (d) frequency of rheme to theme relationships. Each of these four linguistic features, including how they were operationalized in our analysis, is described in more detail in the relevant section of the findings below.

Second, from the perspective of our work on the LISELL project (Buxton et al., 2013) we also consider the language of science to include the specific linguistic constructions used to express the scientific investigation practices of the LISELL pedagogical model, such as the language of controlling variables to design a fair test, the language of coordinating hypothesis, observation, and evidence, and the language of explaining cause-and-effect relationships. These three linguistic frames served as the contexts for the questions, while the four functional linguistic features served as the focus of the analysis. Next, we describe the patterns we found in the student responses when analyzed along the four functional linguistic dimensions.

Findings

Technical Vocabulary Usage
We consider technical vocabulary to include vocabulary that is unique to the context of science (e.g., photosynthesis, insulator) as well as vocabulary that has both an everyday and a science-specific usage if used appropriately in the scientific sense (e.g., matter, conduct). In our analysis of technical vocabulary used by students on the assessment items, we began by counting all uses of technical vocabulary on each focal question. Next, we conducted frequency counts of each technical vocabulary word used in responses to each question. Table 2 shows the three most common technical vocabulary words used in student responses to each question (in decreasing frequency of use) as well as the average number of technical vocabulary words used for each question by each group of students on both pre- and post-assessment.

We noted three patterns in students' uses of technical vocabulary that seemed to be consistent across the different questions, and therefore, across the different investigation practices. First, in every case but one, there was growth in the average number of technical vocabulary words used from the pre-assessment to the post-assessment. In some cases this growth was quite small and in other cases more substantial, but the pattern of increased technical vocabulary at the end of the year was consistent (recall that we are not comparing growth in the same students, but rather averages of randomly selected groups at pre- and post-assessments). At the same time, we noted that the overall usage of technical vocabulary was fairly low, often averaging less than one example per student per question.

Second, we saw a pattern in technical vocabulary usage related to students' linguistic choices on the test. While on the pre-assessment, students in all three language groups showed very similar levels of technical vocabulary usage in their responses, on the post-assessment, students who read the questions in both languages but wrote in English consistently used the most technical vocabulary, and as a group, showed the most growth in technical vocabulary.

Third, we noted a difference in the nature of the technical vocabulary words used most frequently between pre-assessment and post-assessment. On the pre-assessment, the most common technical vocabulary words used were those related to the science content of the question, for example, words related to electricity for the question about electrical circuits or words related to plant growth for the question about the effect of light on plants. In contrast, on the post-test, the most frequently used technical vocabulary words were more often vocabulary related to science investigation practices, for example, hypothesis, effect, or observation.

Table 2 *Technical Vocabulary Usage*

Question topic & form	Written in Spanish, pretest	Written in Spanish, posttest	Read in Both, pretest	Read in Both, posttest	Read in English only, pretest	Read in English only, posttest
Variables Blue	Experiment growth design (.8)	Experiment independent variable (1.1)	Soil growth fertilizer (.7)	Soil, experiment, independent (1.3)	Growth soil data (.6)	Variable independent experiment (.9)
Variables Pink	Variable strength experiment (.5)	Experiment effect control (.3)	Variable experiment effect (.3)	Depend, variable, experiment (.7)	Variable experiment independent (.5)	Variable depend experiment (.6)
Hypothesis Blue	Attract temperature thermometer (.6)	Temperature observe hypothesis (.8)	Attract reflect temperature (.4)	Hypothesis observation temperature (1.0)	Attract reflect absorb (.6)	Hypothesis absorb observation (.8)
Hypothesis Pink	Temperature hypothesis dissolve (.7)	Temperature hypothesis observe (.9)	Temperature hypothesis evaporate (.4)	Hypothesis temperature observe (1.0)	Temperature dissolve hypothesis (.5)	Temperature thermometer hypothesis (.6)
Cause & Effect Blue	Electricity conduct circuit (1.0)	Electricity cause conduct (1.4)	Electricity conduct effect (.7)	Energy cause circuit (1.4)	Electricity circuit effect (.8)	Electricity effect circuit (.9)
Cause & Effect Pink	Population increase effect (.4)	Increase population effect (.5)	Decrease disease population (.4)	Effect decrease population (.8)	Population disease increase (.4)	Cause increase disease (.8)

Grammatical Drift Toward More Stable Linguistic Classes

As noted in the earlier discussion of SFL, the language of science uses grammatical metaphor that results in linguistic drift from less stable linguistic classes, such as circumstances and processes, toward more stable linguistic classes, such as entities and qualities that have increased persistence through time and are thus easier to analyze and categorize. *Nominalizations*, in which circumstances or processes are turned into abstract nouns, are the most common type of grammatical drift in scientific texts. Nominalizations also serve to remove the actor from the action so that abstract nouns perform most of the work in the clause (e.g., the collection of samples; an analysis of the problem).

In our analysis, we recorded the use of any nominalizations or other grammatical drift from circumstances or processes toward entities or qualities in the student responses. Table 3 shows the percentage of students for each language usage group who used at least one clear example of nominalization or other grammatical drift in their response. Below the table, examples from each of the focus questions are provided.

Table 3 *Usage of Nominalizations and Grammatical Drift*

Question	Spanish, pretest	Spanish, posttest	Both, pretest	Both, posttest	English only, pre	English only, post	***Average***
Variables Blue	33%	67%	20%	40%	20%	20%	***33%***
Variables Pink	53%	47%	33%	33%	33%	27%	***38%***
Hypothesis Blue	40%	60%	33%	33%	20%	20%	***34%***
Hypothesis Pink	67%	67%	27%	33%	13%	7%	***36%***
Cause & Effect Blue	27%	53%	27%	40%	13%	13%	***29%***
Cause & Effect Pink	40%	40%	40%	20%	7%	20%	***28%***
Average	***43%***	***56%***	***30%***	***33%***	***18%***	***18%***	***33%***

Examples of nominalizations and grammatical drift toward qualities and entities:

Variables experiment with plants (blue form)

- The type of plant and the amount of light it gets
- The growth of the plants
- The plant's growth under certain light conditions
- How much brightness is needed to grow

Variables experiment with weight lifting (pink form)

- The weight of the person
- Your body stance may get better over time
- Strength of the lifter
- The duration of the exercise

Hypothesis and evidence experiment with light and dark colors (blue form)

- Dark colors absorb more heat
- A black bucket would heat the most after sitting in the sun
- Light colored clothing doesn't heat a person as fast
- Black color heats more than white color

Hypothesis and evidence experiment with salt and boiling water (pink form)

- After boiling salted water is warmer than unsalted water
- Salt content lowers the temperature
- The effect of salt on boiling water would be the same
- Water without salt does not heat as much as water with salt

Cause and effect with electric circuits (blue form)

- The rubber does not conduct electricity
- The nail, since it's metal, will pass the electricity

- Rubber is not a conductor so no energy is passed
- The nail is a conductor and lets electricity pass through

Cause and effect with ecosystems (pink form)

- The effect is large fishes eat small fishes
- The sickness struck the small fishes
- The disease causes a decrease in population
- The effect is a reduced amount of large fish

Again, several patterns are visible in how students made use of grammatical drift in their responses. First, students who wrote at least some of their responses in Spanish were much more likely to use grammatical drift toward entities and qualities in their writing. Students who wrote in Spanish were also the only group that showed a large increase in examples of grammatical drift on the post-assessment when compared with the pre-assessment. Students who read the questions in both languages, but wrote in English, also showed a greater use of grammatical metaphor at both the start and the end of the year when compared to Latino/a students who read and wrote in English only.

Second, the inquiry focus of the questions seemed to have some influence on the frequency of use of grammatical metaphor. Specifically, the two cause and effect questions resulted in lower usage of grammatical drift than the questions focused on controlling variables and on coordinating hypothesis, observation, and evidence.

Finally, we noted that as with technical vocabulary usage, the overall quantity of grammatical drift toward entities and qualities in these students' writing was fairly low, indicating that while middle school students were beginning to make use of this aspect of the language of science, it remained somewhat foreign to their science writing, which typically had a more conversational language structure.

Lexical Density

Lexical density is a measure of the ratio of content words to grammatical words in any given text (spoken or written). Content words are defined as the words that are critical for describing the content of what is being said or written. These include most nouns (except pronouns), most adjectives (except *pro-form adjectives*, such as "so" or "too," that replace more descriptive adjectives), most verbs (except auxiliary verbs) and most adverbs (except *pro-form adverbs*, such as "here," "how," "why," etc.). Grammatical words (also sometimes called functional words) include pronouns, prepositions, conjunctions, auxiliary verbs (e.g. can, could, will), pro-form adverbs and adjectives, determiners (e.g. "a," "the," "my") and interjections (e.g., "wow"). The average ratio of content words to grammatical words, known as lexical density, is approximately 0.5 or 50% in standard non-technical written text. The ratio for normal spoken conversation is typically less than 50% and for technical or specialized academic writing (such as science texts) it is typically greater than 50%. We calculated the lexical density for each student's response to each focal question by dividing the number of content words in the response by the total number of words and expressing it as a percentage. Table 4 shows the average lexical density (with standard deviation) for each question and for each group of students.

Several patterns can be seen in the lexical density of the student responses. First, there was a small but consistent increase in the lexical density of students' responses between the start and the end of the year. Of the 18 possible comparisons (three language groups each for six questions), 14 of these show an increase between pre- and post-assessment and 9 of these comparisons (half of the total) show an increase of at least 3% in lexical density; a meaningful change, given that lexical density rarely varies outside of the 45%-60% range. Thus, at the end of the year, students' written responses on most questions showed a trend toward language that was less conversational and more academic.

Table 4 *Lexical Density*

Question	Spanish, pretest LD (SD)	Spanish, posttest	Both, pretest	Both, posttest	English only, pretest	English only, posttest	***Average***
Variables Blue	49% (.04)	54% (.06)	53% (.08)	56% (.06)	52% (.08)	55% (.07)	***53%***
Variables Pink	51% (.09)	51% (.05)	54% (.10)	55% (.06)	50% (.08)	51% (.09)	***52%***
Hypothesis & Evidence Blue	51% (.07)	55% (.08)	60% (.07)	62% (.07)	56% (.06)	59% (.08)	***57%***
Hypothesis & Evidence Pink	49% (.08)	53% (.08)	52% (.06)	53% (.07)	47% (.06)	52% (.06)	***51%***
Cause & Effect Blue	51% (.07)	55% (.07)	56% (.09)	56% (.07)	55% (.06)	55% (.07)	***55%***
Cause & Effect Pink	49% (.10)	55% (.05)	57% (.07)	56% (.06)	54% (.07)	54% (.07)	***54%***
Average	***50%***	***54%***	***55%***	***56%***	***52%***	***54%***	***54%***

Examples of more conversational responses with lower lexical density:

- When you have something that's like dark colors it's going to stay hot for a long time.
- If I build something with wires and I use a bulb with it, then if I give it power, the bulb will light up.
- I'll put one of my plants in the closet where it's dark and I'll put my other plant in the window where it can get lots of sunlight, and then I'll check out every day how they're growing and write it down.

Examples of more academic responses with higher lexical density:

- Light color shirts absorb less heat than dark color shirts so you stay cooler.
- The nail conducts electricity and completes the circuit and lights the bulb.
- I will compare how sunlight and a light bulb affect plant growth.

Second, when comparing the student groups based on language choices, we noted that the students who read the questions in both languages but wrote in English exhibited the highest lexical density in their writing, both on the pretest and on the posttest. Students who chose to write in Spanish showed the greatest average increase in lexical density between the pre- and post-assessments. These Spanish language responses began with the most conversational language on the pre-assessments, but on the post-assessments were more typical of the grammar of written language and were comparable to the language used by the Latino/a students who read and wrote in English only.

Rheme to Theme Structure

In systemic functional linguistics, the theme is the main idea of a sentence and the rheme is what the speaker/author is saying about that main idea. In science texts, it is common for the rheme of one sentence to become the theme of the next sentence. In this way, the details regarding a technical idea can be elaborated in one sentence and then assigned a label (technical vocabulary) that can be used in the next and subsequent sentences to represent the idea that was elaborated in the original sentence. This allows technical text to be more concise, by not continually restating lengthy descriptions. In our analysis of student writing on the LISELL assessment, we recorded any examples we saw of students using rheme to theme linkages in their responses. Our initial intention was to present descriptive statistics of rheme to theme patterns in table form as we have done for the other linguistic features, followed by examples. We found, however, that the students' use of rheme to theme linguistic structure was so rare that providing the numerical data did not seem meaningful. While there was no *typical* usage of the rheme to theme structure in these student responses (the typical usage being none), we provide a few examples of what these structures looked like when they were used.

Variables experiment with plants (blue form)

- They replaced the sun with the light/ the light is now inside instead of outside
- Plant's growth will be healthier in sunlight/ Sun is one of the primary sources of energy

Variables experiment with weight lifting (pink form)

- The bar was too light so I added more weight/ extra weight would be harder and get me stronger
- Lifting weights burns fat/ with less fat weighing you down, you become stronger

Hypothesis and evidence experiment with light and dark colors (blue form)

- The temperature rising on both buckets/ the black bucket was hotter than the light bucket
- When it's hot you don't want to wear dark clothes/ dark color absorbs heat

Hypothesis and evidence experiment with salt and boiling water (pink form)

- Water without salt cools quicker than water with salt/ the salt keeps the water warmer
- After boiling the water the pot with salt has a higher temperature/ the higher temperature means the hypothesis is right

Cause and effect with electric circuits (blue form)

- The effect is that it will conduct electricity/ electricity passes through the metal
- The circuit can't pass energy through the rubber band/ the rubber band is not a conductor

Cause and effect with ecosystems (pink form)

- If there's no algae the small fish will die/ if there's no small fish, the big fish will die

- Population of algae will increase/ algae increase because there will be no small fish to eat it

These examples indicate that even in the rare instances when students did use rheme-to-theme structure in their responses, the structure was not used to accomplish its usual linguistic function in scientific discourse: to imbue a term with nuanced meaning and then to reuse that term to encapsulate those nuances. Rather, these repetitions of the object of one phrase as the subject of the next seemed to lack a clear functional purpose in the students' writing.

Discussion

As we noted at the outset of this paper, one goal of the LISELL project was to seek better measures of bilingual students' emergent science learning. In line with the latest reform documents (NRC, 2011, 2009), we view critical thinking and critical communication skills as central to worthwhile science learning. Examining students' written responses to science investigations may be one way for both teachers and researchers to understand how this type of science learning develops in students. Further, we wished to add to the limited research on the potential value of linguistic accommodations for emergent bilingual learners on assessments. A better research base is needed both to help teachers develop more useful formative assessments of students' thinking for use in their classrooms and to inform large-scale assessment design.

The LISELL assessment employed a range of linguistic, cognitive, and visual clarifications intended to help emergent bilingual learners express their ideas as clearly as possible. We believed that analyzing student responses using linguistic categories taken from systemic functional linguistics would allow us to identify response patterns that highlighted any potential benefits of bilingual assessments for emergent bilingual learners. Below, we interpret the findings from our functional linguistic analysis of emergent bilingual students' responses on this assessment.

Technical Vocabulary

In our analysis, technical vocabulary included words that are unique to science, as well as words that have both everyday and science-specific usage. Patterns in technical vocabulary usage in our sample of assessments included (a) growth in the average number of technical vocabulary words used from the pre-assessment to the post-assessment; (b) increased technical vocabulary usage on the post-assessment by students who read the questions in both languages; and (c) a shift in the nature of technical vocabulary used, from a focus on words related to the science content of the question, to an increased focus on vocabulary related to the LISELL science inquiry practices.

The first pattern of increased technical vocabulary from pre- to post-assessment can be explained by the LISELL project's attention to vocabulary as part of our professional learning with teachers. If teachers focus more on the vocabulary needs of their emergent bilingual learners, we would expect students to increase their technical vocabulary usage on the assessment. The second pattern, that students who read the questions in both Spanish and English used more robust technical vocabulary, might be explained by the fact that the LISELL assessment prompts used a number of technical vocabulary terms that were important to the question. Students who read the questions in both languages had double the exposure to these terms as they were considering the question, perhaps centering this vocabulary more clearly in their minds as they were composing their responses. Finally, the pattern that students used more technical vocabulary related to the question scenario at the start of the year, and more vocabulary related to the LISELL investigation practices at the end of the year, may indicate that teachers in the project focused on the importance of science investigation practices, including the language of those practices, throughout the year. It could be hoped that students would reflect that emphasis by making greater use of the language of scientific investigation in their responses at the end of the year.

Grammatical Drift

In our analysis, grammatical drift represented movement along the continuum from less stable linguistic classes, such as circumstances and processes, to more stable linguistic classes such as qualities and entities. Patterns we saw in student responses included the following: (a) students who wrote their responses in Spanish were more likely to use more stable linguistic classes; (b) the inquiry focus of the question seemed to influence the linguistic classes used; (c) the overall linguistic pattern in student responses seemed to favor the less stable linguistic classes.

The first pattern of Spanish language responses using more stable linguistic classes may be at least partly explained by the common use of subjunctive mood in Spanish grammar (in English, the subjunctive mood exists, but is rarely used). The subjunctive mood, which expresses conditionality, seems to facilitate grammatical drift toward qualities and entities. Thus, student responses written in Spanish, or framed mentally in Spanish even if sometimes written in English, seem to be more congruent with this key grammatical structure of the language of science. The second pattern we noted, that the inquiry focus on cause-and-effect relationships seemed to show the lowest amount of grammatical drift, might be at least partially explained by the particular nature of cause and effect questions, in that causes and effects seem more conducive to expression as circumstance and process, rather than as quality or entity. Third, in explaining the overall pattern of students favoring the less stable grammatical classes in their writing, we noted in our earlier discussion of SFL that many middle school students are still fairly new to content-specific technical grammar and remain more comfortable writing as they speak, in conversational language that privileges circumstances and processes. It is for this very reason that careful attention to the linguistic structures of the language of science is a central component of the LISELL project. We expect that students, with more practice, will begin to demonstrate a shift toward using more stable qualities and entities in their writing about science topics.

Lexical Density

In our analysis, we considered lexical density to be the ratio of content words (needed to describe the substance of what is being said) to grammatical words (which function to support and clarify the content). Patterns we saw in the lexical density of student responses included the following: a) consistent increase in the lexical density of students' responses between the start and the end of the year, and b) students who read the questions in both languages had the highest lexical density in their writing, while students who wrote in Spanish showed the greatest average increase from the start to the end of the school year.

The first pattern, a general increase in lexical density over the school year, may be at least partially explained by the LISELL project focus on scientific investigation practices and the specific language used to express those practices. Because lower lexical density is often a result of speech patterns that are typical of less precise spoken conversations ("well maybe I think that"), practice engaging with specific investigation practices that require more exact language (e.g., the language of variables, the language of hypothesis) might be expected to result in increased lexical density. The second pattern, in which students who engaged with the bilingual features of the assessment exhibited higher lexical density than those Latino/a students who did not engage with the Spanish language support, may indicate that reading and thinking about the questions in two languages gave students extra models of what the language of science should sound like in the context of the specific questions being asked. Students could then use those models when framing their responses.

Rheme to Theme Structure

Finally, rheme-to-theme structure in our analysis refers to a common pattern in scientific texts, in which the rheme of one sentence (what is being said about the main idea) becomes the theme (or main idea) of a subsequent sentence, thus building on and elaborating complex ideas. While we expected to see a similar increase in this linguistic structure over the course of the year, as we did for the other linguistic structures we analyzed, in fact, rheme-to-theme structures persisted in being quite rare in the student responses.

One possible explanation may be the structure of our assessment response format, which used charts to frame student responses (see Appendix). These charts may not be conducive to connecting ideas from one part of the question to the next, as students might think of each column in the chart as a separate and unrelated portion of the question. Additionally, many students chose to write only one-sentence responses to questions, meaning that there was no chance to connect rhemes and themes across sentences. Similarly, many students who did write more extensively used grammatically incorrect run-on sentences rather than multiple, distinct sentences. More separate sentences would likely have produced more obvious examples of rheme-to-theme structures. Thus, while students' use of technical vocabulary, grammatical drift, and lexical density all seemed to be influenced by their participation in the LISELL project, students were not similarly influenced when it came to adopting the use of rheme-to-theme structures in their science writing.

Implications

Implications for Teachers' Classroom Practices and Formative Assessment

A great deal of attention is currently placed on formative assessments as a way to both guide teachers' classroom practices and to prepare students for high-stakes standardized assessments. Special attention is often given to the preparation of emergent bilingual students, as well as students from other groups that are perceived to be at risk on test performance (Fang & Schleppegrell, 2008). An insufficient amount of attention, however, is paid to the general academic language and content specific language that is fundamental to demonstrating what one knows on any assessment (Solano-Flores, 2008).

Teacher-created formative assessments, such as constructed response written questions, can serve multiple purposes that benefit both students and teachers. First, students can more readily engage in the kinds of critical thinking processes, such as explaining cause-and-effect relationships or coordinating hypothesis and evidence, that are essential for inquiry-based learning, both in and beyond science. Second, students need to make their own language choices regarding how to express their understanding, thus increasing the linguistic challenge but also the language-learning opportunities of the task. Third, teachers can gain much clearer insight into how each student is making sense of the content ideas as well as how each student is able to communicate that thinking through writing. Teachers should be able to use this information to make more informed instructional decisions. As the findings from this study indicate, gaining mastery of the language of science is challenging but attainable for emergent bilingual learners in middle school, given ongoing support and modeling. This support must come primarily from teachers, but well-crafted assessments can facilitate this support as well.

Implications for Large-Scale Standardized Assessments

Designers of large-scale standardized assessments strive to create the conditions for gathering reliable and valid measures of student performance and progress over time. In the current educational climate in the U.S. and many other nations, a great deal of faith is placed in the ability of these assessments to provide accurate data for decision-making with serious consequences for students (e.g., grade retention), teachers (e.g., merit pay), and schools (e.g., school report cards and public opinion). There are, however, many possible threats to the ability of assessments to live up to these expectations, including serious questions about the validity of such data as representative of the science knowledge and skills for the growing number of emergent bilingual learners in our classrooms.

Given these realities, it is perhaps somewhat surprising that there is not more research and development to improve assessment accommodations for emergent bilingual learners. As we have noted earlier, the limited research on assessment accommodations for English learners is ambiguous, and the research on the value of bilingual assessments is particularly spotty. Given the various psychological and sociological influences, as well as questions of prior academic experience and support in the home language, it should not be surprising that simply providing a bilingual assessment seems to do little to improve the measure of emergent bilingual learners' abilities. Teachers must prepare students to get the most out of such accommodations through their ongoing instructional and assessment practices in the classroom. Only in this way will students be able to make the best use of a range of assessment accommodations when it comes time to perform on standardized assessments. We should note that the patterns for making use of home language resources that we saw in this small study–such as girls being more likely to use home language resources than boys, and younger middle school students being more likely to do so than older students–point to some initial considerations upon which teachers and school administrators may wish to reflect.

Implications for Further Research

Our findings from this foray into the use of bilingual constructed response assessments to support emergent bilingual learners in demonstrating their knowledge of science inquiry practices through writing the language of science can be characterized as largely positive but with substantive challenges. Our systemic functional linguistic analysis of the written responses of 180 Latino/a students on items from our project-designed written assessment points to a number of areas where emergent bilingual learners demonstrated substantive proficiency and growth using key aspects of the language of science. This seemed to be particularly true for students who made use of the Spanish language resources provided on our assessment. At the same time, the majority of the students assessed still had a long way to go before mastering the language of science for the purpose of explaining their scientific thinking. Our bilingual assessment, as well as the ongoing teacher professional learning that accompanied it, seemed to offer promise, while pointing to the importance of additional research and development work needed to help educators get the most out of formative assessments and classroom writing more generally, with all students, and particularly with emergent bilingual learners.

First, we need more research that not only closely tracks and documents the choices that students make when presented with assessment accommodations such as bilingual assessments but also attempts to understand why students make those choices. Thus, student interviews and think aloud studies are needed, in which students explain their thinking about why and how they do and do not make use of certain accommodations on an assessment. Second, it seems clear that assessment accommodations are not particularly useful without classroom support that scaffolds students' familiarity with and use of those accommodations. Thus, we need more research that connects what teachers and students do together in the classroom while thinking about challenging academic questions with what students do alone in the context of assessments (whether teacher-constructed or standardized). Finally, our study seems to indicate that the analytic framework of systemic functional linguistics can provide a useful lens for understanding how students gradually learn to take ownership of the language of science and use it in nuanced ways. Thus, additional research is needed that applies SFL frameworks to better understand how all students, but especially emergent bilingual learners, may gradually become more familiar and comfortable with the unique linguistic features of the language of science though repeated exposure and practice, both in the science classroom and on science assessments. The results of this study add to our belief that multilingualism is a cognitive as well as a linguistic and social resource that can be leveraged with practice to enhance academic performance such as critical thinking and critical communication skills. More research is needed into how teachers can support multilingual practices across the content areas and how students can better use such practices to demonstrate their emergent content understanding and thinking skills, both during classroom instruction and on assessments.

References

Buxton, C., Allexsaht-Snider, M., & Rivera, C. (2012). Science, language and families: Constructing a model of language-rich science inquiry. In J. Bianchini, V. Atkerson, A. Calebrese Barton, O. Lee, & A. Rodriguez (Eds.). *Moving the equity agenda forward: Equity research, practice and policy in science education.* New York: Springer, 241 – 259.

Buxton, C., Allexsaht-Snider, M., Suriel, R., Kayumova, S., Choi, Y., Bouton, B., & Land, M. (2013). Using Educative Assessments to Support Science Teaching for Middle School English Language Learners. *Journal of Science Teacher Education, 24*(2), 347 – 366.

Chomsky, N. (2004). *On nature and language.* Cambridge, UK: Cambridge University Press.

Duff, P. (2011). Language socialization into academic discourse communities. *Annual Review of Applied Linguistics, 30*, 169 – 192.

Duschl, R., & Grandy, R. (2008). *Teaching scientific inquiry: Recommendations for research and implementation.* Rotterdam, Netherlands: Sense Publishers.

Eggins, S. (2004). *Introduction to systemic functional linguistics.* London, UK: Continuum.

Fang, Z., & Schleppegrell, M. (2008). *Reading in secondary content areas: A language-based pedagogy.* Ann Arbor, MI: University of Michigan Press.

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education*, *92*(5), 799 – 824.

Gunel, M., Hand, B., & Prain, V. (2007). Writing for learning in science: A secondary analysis of six studies. *International Journal of Science and Mathematics Education, 5*(4), 615 – 637.

Halliday, M. A. K. (1994). *An introduction to functional grammar.* London, UK: Edward Arnold.

Halliday, M. A. K. (1999). Grammar and the construction of educational knowledge. In R. Berry (Ed.). *Language analysis, description and pedagogy, Chapter 15.* Hong Kong: Lignan University.

Halliday, M. A. K. (2004). *The language of science.* London, UK: Continuum.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*(1), 55 – 88.

Huerta, M., Lara-Alecio, R., Tong, F., & Irby, B. J. (2014). Developing and validating a science notebook rubric for fifth-grade non-mainstream students. *International Journal of Science Education, 36*(11), 1849 – 1870.

Kieffer, M. J., Lesaux, N. K., Rivera, M., & Francis, D. J. (2009). Accommodations for English language learners on large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research, 79*(3), 1168 – 1201.

Kuhn, D. (2005). *Education for thinking*. Cambridge, MA: Harvard University Press.

Lee, O., Penfield, R. D., & Buxton, C. (2011). Relationship between "form" and "content" in science writing among English language learners. *Teachers College Record, 113*(7), 1401 – 1434.

Levine, G. (2011). *Code choice in the language classroom.* Bristol, UK: Multilingual Matters.

National Center for Education Statistics. (2011). *The condition of education 2011* (NCES 2011-033). Washington, DC: U.S. Department of Education.

Next Generation Science Standards. (2013). Washington, DC: Achieve, Inc. Retrieved from http://nextgenscience.org/next-generation-science-standards

National Research Council. (2009). *Learning science in informal environments: People, places and pursuits.* Washington, DC: National Academy Press.

National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting themes, and core ideas.* Washington, DC: National Academy Press.

Rosebery, A. S., & Warren, B. (Eds.). (2008). *Teaching science to English language learners: Building on students' strengths*. NSTA Press.

Shaw, J. M., Bunch, G., & Geaney, E. (2010). Analyzing the language demands facing English learners on science performance assessments: The SALD framework. *Journal of Research on Science Teaching, 47*(8), 909 – 928.

Siegel, M. A. (2007). Striving for equitable classroom assessments for linguistic minorities: Strategies for and effects of revising life science items. *Journal of Research in Science Teaching, 44,* 864 – 881.

Solano-Flores, G. (2008). Who is given tests in what language by whom, when, and where? The need for probabilistic views of language in the testing of English language learners. *Educational Researcher, 37*(4), 189 – 199.

Sotelo-Dynega, M., Ortiz, S., Flanagan, D., & Chaplin, W. (2013). English language proficiency and test performance: An evaluation of bilingual students with the Woodcock-Johnson III tests of cognitive abilities. *Psychology in Schools, 50*(8), 781 – 797.

Wong Fillmore, L. (2000). Loss of family languages: Should educators be concerned? *Theory into Practice, 39*(4), 203 – 210.

Appendix

Sample Assessment Item

2

2. Lifting weights makes people stronger. You wonder if different **types** of weight lifting exercises affect how strong a person gets. You decide to design an experiment to find out. Think about the experiment you would design and answer these questions in the table.

a) What is the **Independent Variable** in your experiment about lifting weights? (The independent variable is the variable that you change or manipulate)

b) What is the **Dependent Variable** in your experiment about lifting weights? (The dependent variable is the variable that changes in response to the independent variable)

c) List **2** other **Variables** you will need to **control** in your experiment.

d) Then, use **scientific language** to describe your experiment about lifting weights.

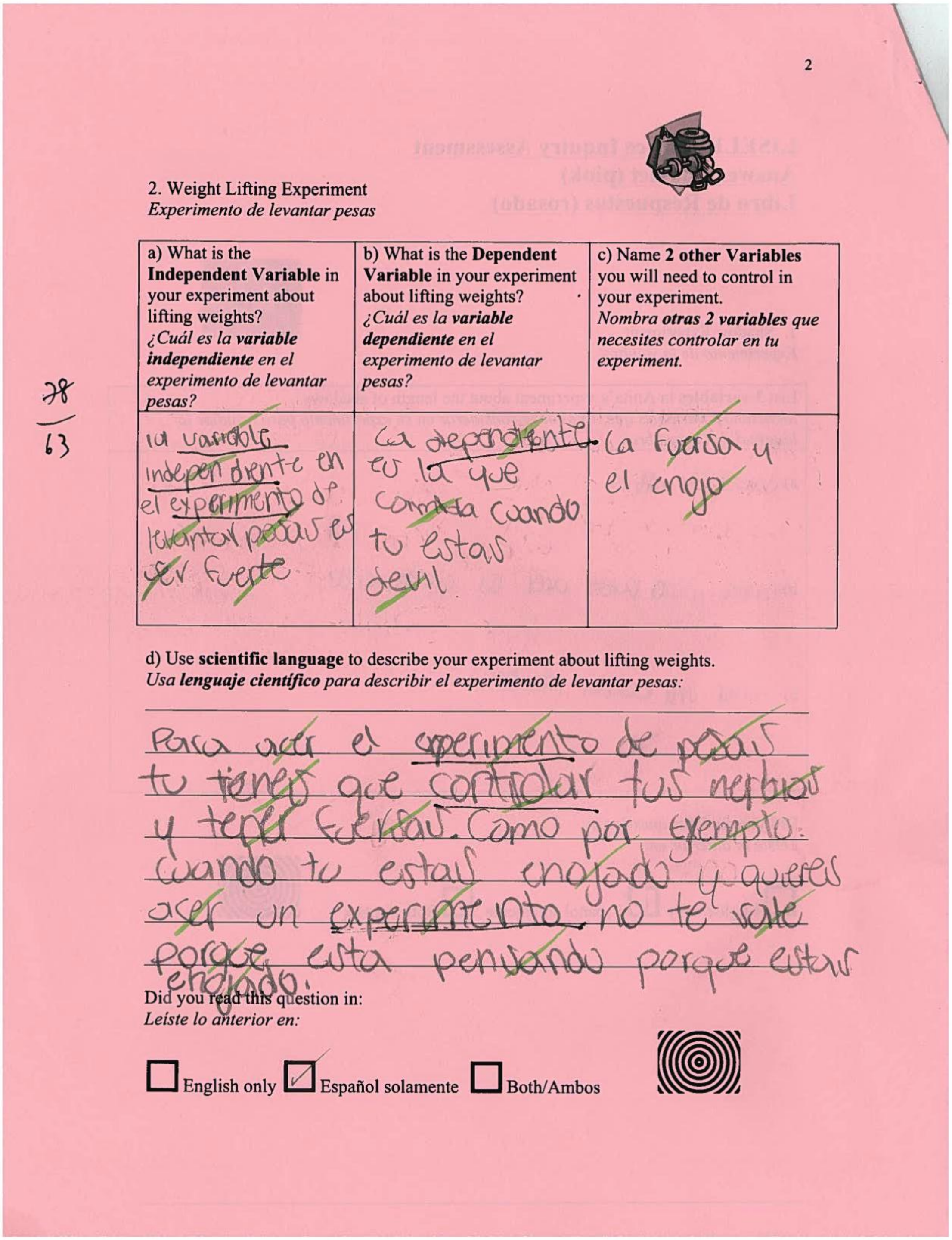
2. Levantar pesas hace a la gente más fuerte. Te preguntas cómo los diferentes ejercicios afectan la fortaleza de las personas. Decides diseñar un experimento para resolver tu duda. Piensa acerca del experimento que diseñarías y después, responde estas preguntas en la tabla.

*a) ¿Cuál es la **Variable Independiente** en tu experimento de levantar pesas? (La variable independiente es la que tú manipulas o cambias)*

*b) ¿Cuál es la **Variable Dependiente** en el experimento de levantar pesas? (La dependiente es la que cambia en cuando la variable independiente cambia)*

*c) Menciona al menos **2 variables** que tendrás que **controlar** en tu experimento.*

*d) Utiliza **lenguaje científico** para describer el experimento de levantar pesas.*



2

2. Weight Lifting Experiment
Experimento de levantar pesas

a) What is the **Independent Variable** in your experiment about lifting weights? *¿Cuál es la **variable independiente** en el experimento de levantar pesas?*	b) What is the **Dependent Variable** in your experiment about lifting weights? *¿Cuál es la **variable dependiente** en el experimento de levantar pesas?*	c) Name **2 other Variables** you will need to control in your experiment. *Nombra **otras 2 variables** que necesites controlar en tu experiment.*
la variable independiente en el experimento de levantar pesas es ser fuerte	La dependiente es la que [illegible] cuando tu estas devil	La fuerza y el enojo

78/63

d) Use **scientific language** to describe your experiment about lifting weights.
*Usa **lenguaje científico** para describir el experimento de levantar pesas:*

Para acer el experimento de pesas tu tienes que controlar tus nerbios y tener fuerzas. Como por ejemplo: cuando tu estas enojado y quieres aser un experimento, no te sale porque esta pensando porque estar enojado.

Did you read this question in:
Leiste lo anterior en:

☐ English only ☑ Español solamente ☐ Both/Ambos