The International Research Foundation
for English Language Education

Title of Project:
Investigating the Potential of Multimodal Tasks to Promote More Equitable Assessment of English Learners in Science

Researcher:
Scott Grapin
sgrapin@miami.edu

Doctoral Institution
New York University

Current Affiliation
University of Miami

Research Supervisor:
Prof. Lorena Llosa
New York University


Scott Grapin

TIRF Research Topic Investigated:
Language Assessment
Content-Based Instruction

Final Report

Motivation for the Research

Across the United States, a fast-growing population of English learners (ELs) and former ELs is learning and being assessed on academic content in English. In the midst of these demographic changes, the most recent wave of academic standards in U.S. K-12 education presents fundamental shifts in the way content areas are taught and learned (Council of Chief State School Officers, 2012). A key feature of the latest standards is their emphasis on disciplinary practices that integrate multiple modalities (e.g., visual, written, oral). For example, students are expected to produce digital texts in English language arts, construct arguments using symbolic representations in math, and develop models using visual representations in science.

The changes in teaching and learning heralded by the standards call for parallel shifts in assessment, including the assessments that teachers regularly implement in their classrooms as part of instruction (Heritage et al., 2015). Traditionally, assessments of content learning have been carried out through written language, with nonlinguistic modalities (e.g., visuals) being deprioritized or excluded altogether (Fernandes et al., 2017). However, with the latest standards emphasizing the multimodal nature of content learning, it is critical that assessments embrace a broader spectrum of linguistic and nonlinguistic modalities. Assessments focused exclusively on written language could miss ways that all students, and ELs in particular, demonstrate their content learning using their full range of meaning-making resources.

The purpose of this study was to investigate the extent to which *multimodal assessment tasks* (i.e., tasks that elicit responses in multiple modalities as part of their design) provide

information about students' content learning that traditional written language assessments may overlook and whether this is particularly the case for ELs. Specifically, this study was an initial inquiry into one type of multimodal task focused on the science disciplinary practice of "developing and using models" in the Next Generation Science Standards (NGSS).

Research Questions

The study was guided by the following research questions (RQs):

1. How does performance on science modeling tasks in the *visual modality* compare with performance in the *written modality* for students at varying levels of English proficiency?
2. How does eliciting responses to the same tasks in the *oral modality* provide additional information about students' science understanding beyond visual and written modalities?

Research Methodology

Context and Participants

Participants were 393 fifth-grade students in the science classrooms of eight teachers who implemented the Science And Integrated Language (SAIL) curriculum (Lee et al., 2019). More than half of the students were non-ELs (61%), while approximately one third were former ELs (33%) and a much smaller percentage (6%) were current ELs. A subsample of 35 students from all three EL classifications (10 non-ELs, 12 former ELs, 13 current ELs) participated in oral interviews.

Science Modeling Tasks

Four science modeling tasks–one from each end-of-unit assessment in the curriculum–were used in this study. The four tasks addressed NGSS performance expectations in physical science, life science, Earth science, and space science, respectively. In the first part of each task, students responded using drawings and symbols (i.e., visual modality). In the second part, students responded using written language (i.e., written modality). Student responses were scored separately in visual and written modalities using detailed science scoring criteria. After each end-of-unit assessment, the researcher met individually with each student interview participant. Students were provided a copy of their visual response and asked to respond to the same prompt in the oral modality.

Data Analysis

To compare student performance in visual versus written modalities (RQ1), visual and written scores were cross-tabulated for each task. Next, a two-way repeated measures analysis of variance (RMANOVA), with modality (visual, written) as the within-subjects factor and EL classification (non-EL, former EL, current EL) as the between-subjects factor, was conducted for each task to test for significant differences in performance between the modalities (i.e., main effect for modality) and whether these differences varied significantly by EL classification (i.e., interaction effect between modality and EL classification). To investigate how eliciting responses in the oral modality provided additional information beyond visual and written modalities (RQ2), interviews were transcribed using multimodal conventions (Flewitt et al., 2017) and coded inductively (Strauss & Corbin, 1998).

Summary of Findings

RQ1

On all four tasks, approximately half of the students performed differently in visual versus written modalities. For these students, the interpretation of their science understanding was different depending on the modality of assessment. Consistent with multimodal theory and research (e.g., Jewitt et al., 2001), these differences were attributable, in part, to the unique affordances of the visual modality as compared to the written. For example, the visual modality, governed by "the logic of display in space" (Kress, 2000, p. 339), elicited students' ideas about the relative position and movement of Earth in the space science task.

However, contrary to what was hypothesized, current ELs did not perform consistently better in the visual modality as compared to the written. In the two tasks that involved communicating about abstract nonvisible science processes (physical science and life science), students did not perform better in either modality. In contrast with the conventional wisdom of visuals as accommodations for ELs (see Grapin, 2019 for a critique), the visual modality posed discipline-specific representational demands (e.g., knowing to use dots to represent gas particles and arrows to represent energy transfer). Thus, while a hypothesis motivating the study was that ELs would perform better in the visual because they were emergent users of language, the findings revealed that all students, including ELs, were emergent representers.

Nonetheless, there was evidence that responding visually was beneficial when the representational demands were not as high. In the two tasks that involved communicating about concrete visible science processes (space science and Earth science), all students performed better in the visual modality than in the written. Furthermore, there was descriptive evidence in one task (space science) that responding visually was particularly beneficial to current ELs, who were able to compensate for imprecise written responses with accurate representations of Earth's orbit in their visual responses. Thus, the assumption underlying multimodal accommodations, namely, that ELs will be able to communicate their understanding better in nonlinguistic modalities than through language alone, was partially supported by this study, but the findings suggest that whether such accommodations are effective may depend crucially on the nature of the content being assessed (e.g., whether it involves concrete vs. abstract representation).

RQ2

Beyond visual and written modalities, eliciting responses in a third modality–the oral modality–provided additional information about students' science understanding. While many current ELs and former ELs used technical science terms inaccurately in their written responses (e.g., mixing up "orbit" and "rotate"), these students used gesture in combination with everyday language to demonstrate sophisticated understanding in the oral modality. In contrast, some non-ELs appeared to use technical terms accurately in their written responses, but their oral responses called into question whether they understood the underlying science ideas. Thus, multimodal assessment emerged as useful for "teasing apart [content] knowledge from language used to display that knowledge" (National Academies of Sciences, Engineering, and Medicine, 2018, p. 221). In assessments that restrict their focus to the written modality, ELs who use disciplinary language less accurately or less frequently may be penalized despite having well-developed content understanding, whereas non-ELs may be judged proficient with content regardless of whether their language is underpinned by content understanding.

Finally, it was not only the affordances of the oral modality by virtue of being oral but also the affordances of dynamic interaction in this modality that allowed current ELs and former

ELs to go beyond what they could do independently. While dynamic assessment has been applied extensively to research in foreign language classrooms (e.g., Lantolf & Poehner, 2010), this study found that providing mediation in the oral modality (e.g., asking students, "What do you mean?" and "What's 'it'?") was effective for eliciting students' science ideas while also supporting ELs with the language to communicate those ideas more precisely and explicitly.

Implications

Overall, this study highlights the potential of multimodal assessment for reframing a *deficit* view of ELs perpetuated by a narrow focus on their independent written language performance into an *asset* view that makes visible the expansive meaning-making resources that ELs bring to content learning. For teachers of ELs in the content areas, the findings point to the importance of gathering evidence of learning in multiple modalities and triangulating across those modalities to arrive at a coherent interpretation of what students know and can do. Importantly, teachers must guard against the (largely intuitive) assumption that visuals are *the* answer to the content assessment of ELs—an assumption that is particularly tenuous when dealing with abstract content ideas and discipline-specific norms for communicating those ideas. In a similar manner, teachers must guard against the belief that disciplinary language necessarily equates to more sophisticated content understanding, particularly when dealing with technical terms in the content areas that densely pack sophisticated ideas and therefore can mask a lack of content understanding. When visual representation is more abstract and technical terms are involved, assessment may require a greater degree of triangulation (e.g., asking students to explain orally what they represented visually or expressed in writing) to diagnose specific sources of difficulty (e.g., representational, linguistic, and/or conceptual). By providing more accurate and granular information about each student's learning status, multimodal assessment could allow teachers to engage in more contingent assessment with ELs in content classrooms.

In the context of the latest standards in U.S. K-12 education, this dissertation challenges three assumptions about what "counts" as evidence of content learning that have permeated the assessment of ELs in the content areas. First, by highlighting the indispensable contribution of nonlinguistic modalities to disciplinary meaning-making, this dissertation challenges traditional assumptions of content learning that privilege linguistic modalities and written language in particular. Second, by highlighting how everyday language can be a resource for communicating disciplinary ideas, it challenges traditional assumptions of language use in the content areas that privilege disciplinary language and *the* language of science. Third, by highlighting how a complete picture of learners' abilities cannot be captured by their independent performance alone, it challenges traditional assumptions of assessment that privilege standardization and reliability over contingency and responsiveness. By expanding what "counts" as evidence of content learning beyond traditionally privileged forms of expression, this dissertation presents a challenge to the field to (re)envision the equitable assessment of ELs in content classrooms.

References

Abedi, J., & Herman, J. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record, 112*(3), 723-746.

Abedi, J., Hofstetter, C. H., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Educational Research, 74*(1), 1-28.

Achieve. (2019). *Science assessment task screening tools*. https://www.nextgenscience.org/taskscreener

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn science. *Science, 333*(6046), 1096-1097.

Alonzo, A., & Gotwals (Eds.). (2012). *Learning progressions in science: Current challenges and future directions*. Sense.

Arnaut, K., Blommaert, J., Rampton, B., & Spotti, M. (Eds.). (2016). *Language and superdiversity*. Routledge.

Avenia-Tapper, B., & Isacoff, N. M. (2015). Explicitness in science discourse: A Gricean account of income-related differences. *Language and Education, 30*(1), 58-71.

Bachman, L. (2005). Building and supporting a case for test use. *Language Assessment Quarterly, 2*(1), 1-34.

Bachman, L., & Palmer, A. (1996). *Language testing in practice.* Oxford University Press.

Bachman, L., & Palmer, A. (2010). *Language assessment in practice.* Oxford University Press.

Bailey, A. L. (2007). *The language demands of school: Putting academic English to the test*. Yale University Press.

Bailey, A. L., & Heritage, M. (2014). The role of language learning progressions in improved instruction and assessment of English language learners. *TESOL Quarterly, 48*(3), 480-506.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1986). Is an atom of copper malleable? *Journal of Chemical Education, 63*, 64-66.

Berland, L., Schwarz, C., Krist, C., Kenyon, L., Lo, A., & Reiser, B. (2016). Epistemologies in practices: Making scientific practices meaningful for students. *Journal of Research in Science Teaching, 53*(7), 1082-1112.

Bernstein, B. (1971). *Class, codes and control*. Routledge.

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social-semiotic account of designs for learning. *Written Communication, 25*(2), 166-195.

Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: A social semiotic frame*. Routledge.

Bezemer, J., & Mavers, D. (2011). Multimodal transcription as academic practice: A social semiotic perspective. *International Journal of Social Research Methodology, 14*(3), 191-206.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Block, D. (2014). Moving beyond "lingualism": Multilingual embodiment and multimodality in SLA. In S. May (Ed.), *The multilingual turn: Implications for SLA, TESOL, and bilingual education* (pp. 54-77). Routledge.

Box, G. E., & Draper, N. R. (1987). *Empirical model-building and response surfaces*. Wiley.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge University Press.

Budoff, M., & Friedman, M. (1964). "Learning potential" as an assessment approach to the adolescent mentally retarded. *Journal of Consulting Psychology, 28*(5), 434-439.

Bunch, G. (2014). The language of ideas and the language of display: Reconceptualizing "academic language" in linguistically diverse classrooms. *International Multilingual Research Journal, 8*(1), 70-86.

Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bullet, 56*, 81-105.

Canagarajah, S. (2018). Translingual practice as spatial repertoires: Expanding the paradigm beyond structuralist orientations. *Applied Linguistics, 39*(1), 31-54.

Chamot, A. U., & O'Malley, J. M. (1987). The cognitive academic language learning approach: A bridge to the mainstream. *TESOL Quarterly, 21*(2), 227-249.

Choi, I., Wolf, M. K., Pooler, E., Sova, L., & Faulkner-Bond, M. (2019). Investigating the benefits of scaffolding in assessments of young English learners: A case for scaffolded retell tasks. *Language Assessment Quarterly, 16*(2), 161-179.

Cizek, G. (2010). An introduction to formative assessment: History, characteristics, and challenges. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 3-17). Routledge.

Coffey, J., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching, 48*(10), 1109-1136.

Council of Chief State School Officers. (2012). *Framework for English language proficiency development standards corresponding to the Common Core State Standards and the Next Generation Science Standards.* https://www.ccsso.org/sites/default/files/2017-11/ELPD%20Framework%20Booklet-Final%20for%20web.pdf

Council of Chief State School Officers. (2014). *English language proficiency (ELP) standards with correspondences to K-12 English language arts (ELA), mathematics, and science practices, K-12 ELA standards, and 6-12 literacy standards*. http://elpa21.org/sites/default/files/Final%204_30%20ELPA21%20Standards_1.pdf

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222-251.

Cummins, J., & Early, M. (Eds.). (2011). *Identity texts: The collaborative creation of power in multilingual schools*. Trentham Books.

Cummins, J., Hu, S., Markus, P., & Montero, M. K. (2015). Identity texts and academic achievement: Connecting the dots in multilingual school contexts. *TESOL Quarterly, 49*(3), 555-581.

Dalton, B., Robinson, K. H., Lovvorn, J. F., Smith, B., Alvey, T., Mo, E., . . . Proctor, C. P. (2015). Fifth-grade students' digital retellings and the Common Core: Modal use and design intentionality. *The Elementary School Journal, 115*(4), 548-569.

Danielsson, K. (2016). Modes and meaning in the classroom: The role of different semiotic resources to convey meaning in science classrooms. *Linguistics and Education, 35*, 88-99.

De Backer, F., Van Avermaet, P., & Slembrouck, S. (2017). Schools as laboratories for exploring multilingual assessment policies and practices. *Language and Education, 31*(3), 217-230.

DeBarger A. H., Penuel, W. R., Harris, C. J., & Kennedy, C. A. (2016). Building an assessment argument to design and use next generation science assessments in efficacy studies of curriculum interventions. *American Journal of Evaluation, 37*(2), 174-192.

diSessa, A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction, 22*(3), 293-331.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science: Research into children's ideas.* Routledge.

Duschl, R. A., & Gitomer, D. H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment, 4*(1), 37-73.

Early, M., Kendrick, M., & Potts, D. (2015). Multimodality: Out from the margins of English language teaching. *TESOL Quarterly, 49*(3), 447-460.

Echevarría, J., Vogt, M., & Short, D. J. (2012). *Making content comprehensible for English language learners: The SIOP model.* Allyn & Bacon.

Eder, D., & Fingerson, L. (2001). Interviewing children and adolescents. In J. Holstein & F. Gubrium (Eds.), *Inside interviewing: New lenses, new concerns* (pp. 33-55). Sage.

Erickan, K., & Pellegrino, J. W. (Eds.). (2017). *Validation of score meaning for the next generation of assessments*. Routledge.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* MIT Press.

Fang, Z. (2006). The language demands of science reading in middle school. *International Journal of Science Education, 28*(5), 491-520.

Fernandes, A., Kahn, L. H., & Civil, M. (2017). A closer look at bilingual students' use of multimodality in the context of an area comparison problem from a large-scale assessment. *Educational Studies in Mathematics, 95*(3), 263-282.

Flewitt, R., Hampel, R., Hauck, M., & Lancaster, L. (2017). What are multimodal data and transcription? In C. Jewitt (Ed.), *Handbook of multimodal analysis* (pp. 44-59). Routledge.

Flores, N. (2020). From academic language architecture: Challenging raciolinguistic ideologies in research and practice. *Theory Into Practice, 59*(1), 22-31.

Forbes, C. T., Zangori, L., & Schwarz, C. (2015). Empirical validation of integrated learning performances for hydrologic phenomena: 3rd-grade students' model-driven explanation-construction. *Journal of Research in Science Teaching, 52*(7), 895-921.

Furtak, E., Ruiz-Primo, M. A., & Bakeman, R. (2017). Exploring the utility of sequential analysis in studying informal formative assessment practices. *Educational Measurement: Issues and Practice, 36*(1), 28-38.

García, O., & Li, W. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan.

Giere, R. R. (1988). *Explaining science: A cognitive approach*. University of Chicago Press.

Givry, D., & Roth, W. M. (2006). Toward a new conception of conceptions: Interplay of talk, gestures, and structures in the setting. *Journal of Research in Science Teaching, 43*(10), 1086-1109.

Goldin-Meadow, S. (2004). Gesture's role in the learning process. *Theory Into Practice, 43*(4), 314-321.

Gorin, J. S., & Mislevy, R. J. (2013, September). Inherent measurement challenges in the Next Generation Science Standards for both formative and summative assessment. Paper presented at the Invitational Research Symposium on Science Assessment, Washington, DC. https://www.ets.org/Media/Research/pdf/gorin-mislevy.pdf

Gotwals, A. W., & Songer, N. B. (2009). Reasoning up and down the food chain: Using an assessment framework to investigate students' middle knowledge. *Science Education, 94*(2), 259-281.

Grapin, S. E. (2019). Multimodality in the new content standards era: Implications for English learners. *TESOL Quarterly, 53*(1), 30-55.

Grapin, S. E., Haas, A., Goggins, M., Llosa, L., & Lee, O. (2019). Beyond general-purpose talk moves: Using discipline-specific probes with English learners in the science classroom. *Science and Children, 57*(4), 36-43.

Grapin, S. E., & Llosa, L. (2020). Toward an integrative framework for understanding multimodal L2 writing in the content areas. *Journal of Second Language Writing, 47*, 1-8.

Grapin, S. E., Llosa, L., Haas, A., Goggins, M., & Lee, O. (2019). Precision: Toward a meaning-centered view of language use with English learners in the content areas. *Linguistics and Education, 50,* 71-83.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Guzman-Orth, D., Lopez, A., & Tolentino, F. (2019). Exploring the use of a dual language assessment task to assess young English learners. *Language Assessment Quarterly, 16*(4-5), 447-463

Hakuta, K., & Santos, M. (2012, April). *Understanding language: Challenges and opportunities for language learning in the context of the Common Core State Standards and Next Generation Science Standards.* Paper presented at the Understanding Language Conference, Stanford, CA. http://ell.stanford.edu/papers

Halliday, M. A. K. (1978). *Language as social semiotic*. Edward Arnold.

Halliday, M. A. K., & Mathiessen, C. (2004). *An introduction to functional grammar* (3rd ed.). Edward Arnold.

Haneda, M. (2014). From academic language to academic communication: Building on English learners' resources. *Linguistics and Education, 26,* 126-135.

Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & DeBarger, A. H. (2019). Designing knowledge-in-use assessments to promote deeper learning. *Educational Measurement: Issues and Practice, 38*(2), 53-67.

Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & McElhaney, K. W. (2016). *Constructing assessment tasks that blend disciplinary core ideas, crosscutting concepts, and science practices for classroom formative applications.* SRI International.

Heritage, M., & Heritage, J. (2013). Teacher questioning: The epicenter of instruction and assessment. *Applied Measurement in Education, 26*(3), 176-190.

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice, 28*(3), 24-31.

Heritage, M., Walqui, A., & Linquanti, R. (2015). *English language learners and the new standards: Developing language, content knowledge, and analytical practices in the classroom*. Harvard Education Press.

Hirvela, A., & Belcher, D. (2016). Reading/writing and speaking/writing connections: The advantages of a multimodal pedagogy. In P. K. Matsuda & R. Manchón (Eds.), *Handbook of second and foreign language writing* (pp. 587-612). De Gruyter Mouton.

Hokayem, H., & Gotwals, A. W. (2016). Early elementary students' understanding of complex ecosystems: A learning progression approach. *Journal of Research in Science Teaching, 53*(10), 1524-1545.

Hubber, P., & Tytler, R. (2013). Models and learning science. In R. Tytler, V. Prain, P. Hubber, & B. Waldrip (Eds.), *Constructing representations to learn in science* (pp. 109-133). Sense.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings*. Penguin.

Jaipal, K. (2009). Meaning making through multiple modalities in a biology classroom: A multimodal semiotics discourse analysis. *Science Education, 94*(1), 48-72.

Jewitt, C. (2017). An introduction to multimodality. In C. Jewitt (Ed.), *Handbook of multimodal analysis* (pp. 15-30). Routledge.

Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.

Jewitt, C., Kress, G., Ogborn, J., & Tsatsarelis, C. (2001). Exploring learning through visual, actional and linguistic communication: The multimodal environment of a science classroom. *Educational Review, 53*(1), 5-18.

Johnson, P. (1998). Progression in children's understanding of a "basic" particle theory: A longitudinal study. *International Journal of Science Education, 20*(4), 393-412.

Kane, M. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (pp. 17-64). American Council on Education and Praeger.

Kieffer, M. J., & Thompson, K. D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher, 47*(6), 391-398.

Kopriva, R. J. (2014). *Technology-interactive classroom-embedded modules for measuring challenging math and science skills of ELs*. University of Wisconsin.

Kopriva, R. J., Gabel, D., & Cameron, C. (2009). *Overview of results from the ONPAR elementary and middle school science experimental study with ELs and non-ELs: A promising new approach for measuring complex content knowledge of English learners with lower proficiency*. http://iiassessment.wceruw.org/research/researchPapers/Overview%20Results%20ONPAR%20El%20and%20MS%20Science_7%2012c_13.pdf

Kopriva, R. J., & Wright, L. (2017). Score processes in assessing academic content of non-native speakers: Literature review and ONPAR summary. In K. Ercikan & J. W. Pellegrino (Eds.), *Validation of score meaning for the next generation of assessments: The use of response processes* (pp. 100-112). Routledge.

Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International, 23*(1), 112-127.

Kress, G. (2000). Multimodality: Challenges to thinking about language. *TESOL Quarterly, 34*, 337-340.

Kress, G. (2003). *Literacy in the new media age*. Routledge.

Kress, G. (2009). Assessment in the perspective of a social semiotic theory of multimodal teaching and learning. In C. Wyatt-Smith & J. Cumming (Eds.), *Educational assessment in the 21st century: Connecting theory and practice* (pp. 19-41). Springer.

Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2014). *Multimodal teaching and learning: The rhetorics of the science classroom* (2nd ed.). Bloomsbury Academic.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Lantolf, J. P., & Poehner, M. E. (2010). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research, 15*(1), 11-33.

Larsen-Freeman, D. (2007). Reflecting on the cognitive-social debate in second language acquisition. *The Modern Language Journal, 7*, 773-787.

Lee, O. (2017). Common Core State Standards for ELA/literacy and Next Generation Science Standards: Convergences and discrepancies using argument as an example. *Educational Researcher, 46*(2), 90-102.

Lee, O. (2018). English language proficiency standards aligned with content standards. *Educational Researcher, 47*(5), 317-327.

Lee, O. (2019). Aligning English language proficiency standards with content standards: Shared opportunity and responsibility across English learner education and content areas. *Educational Researcher, 48*(8), 534-542.

Lee, O., Eichinger, C. D., Anderson, W. C., Berkheimer, D. G., & Blakeslee, T. D. (1993). Changing middle school students' conceptions of matter and molecules. *Journal of Research in Science Teaching, 30*(2), 249-270.

Lee, O., & Fradd, S. H. (1998). Science for all, including students from non-English-language backgrounds. *Educational Researcher, 27*(4), 12-21.

Lee, O., & Llosa, L. (2015-2019). *Development of language-focused three-dimensional science instructional materials to support English language learners in fifth grade* [Research project]. Funded by the National Science Foundation [DRK-12 1502507].

Lee, O., Llosa, L., Grapin, S. E., Haas, A., & Goggins, M. (2019). Science and language integration with English learners: A conceptual framework guiding instructional materials development. *Science Education, 103*(2), 317-337.

Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English language arts and mathematics. *Educational Researcher, 42*(4), 223-233.

Lehrer, R., & Schauble, L. (2015). The development of scientific thinking. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (7th ed., pp. 371-388). Cambridge University Press.

Lemke, J. (1998). Multimedia literacy demands of the scientific curriculum. *Linguistics and Education, 10*(3), 247-271.

Lemke, J. (2002). Travels in hypermodality. *Visual Communication, 1,* 299-325.

Li, M., & Storch, N. (2017). Second language writing in the age of CMC: Affordances, multimodality, and collaboration. *Journal of Second Language Writing, 36,* 1-5.

Llosa, L. (2007). Validating a standards-based classroom assessment of English proficiency: A multitrait-multimethod approach. *Language Testing, 24*(4), 489-515.

Llosa, L. (2017). Assessing students' content knowledge and language proficiency. In E. Shohamy & I. Or (Eds.), *Encyclopedia of language and education* (Vol. 7, pp. 3-14). Springer International.

Llosa, L., Lee, O., Jiang, F., Haas, A., O'Connor, C., Van Booven, C. D., & Kieffer, M. (2016). Impact of a large-scale science intervention focused on English language learners. *American Educational Research Journal, 53*(2), 395-424.

Lyon, E. G. (2013a). Conceptualizing and exemplifying science teachers' assessment expertise. *International Journal of Science Education, 35*(7), 1208-1229.

Lyon, E. G. (2013b). What about language while equitably assessing science?: Case studies of preservice teachers' evolving expertise. *Teaching and Teacher Education, 32,* 1-11.

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching, 51*(1), 36-76.

MacDonald, R., Cook, H. G., & Miller, E. (2014). *Doing and talking science: A teacher's guide to meaning-making with English learners*. Board of Regents of the University of Wisconsin System.

Márquez, C., Izquierdo, M., & Espinet, M. (2006). Multimodal science teachers' discourse in modeling the water cycle. *Science Education, 90*(2), 202-226.

Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment, 14*(3-4), 160-179.

Maton, K. (2013). Making semantic waves: A key to cumulative knowledge-building. *Linguistics and Education, 24*(1), 8-22.

Michaels, S., & O'Connor, C. (2015). Conceptualizing talk moves as tools: Professional development approaches for academically productive discussion. In L. B. Resnick, C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through talk and dialogue* (pp. 347-362). American Educational Research Association.

McCaffrey, R. J., Duff, K., & Westervelt, H. J. (2000). *Practitioner's guide to evaluating change with neuropsychological assessment instruments.* Kluwer Academic/Plenum Publishers.

McGrail, E., & Behizadeh, N. (2017). K-12 multimodal assessment and interactive audiences: An exploratory analysis of existing frameworks. *Assessing Writing, 31*, 24-38.

Merritt, J., & Krajcik, J. (2013). Learning progression developed to support students building a particle model of matter. In G. Tsaparlis & H. Sevian (Eds.), *Concepts of matter in science education* (pp. 11-45). Springer.

Messick, S. (2000). Consequences of test interpretation and use: The fusion of validity and values in psychological assessment. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment* (pp. 3-20). Kluwer.

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the Next Generation Science Standards. *Journal of Research on Science Teaching, 55*(7), 1053-1075.

Miller-Cochran, S. (2017). Understanding multimodal composing in an L2 writing context. *Journal of Second Language Writing, 38*, 88-89.

Morgan, M. S., & Morrison, M. (1999). *Models as mediators.* Cambridge University Press.

Moschkovich, J. (1999). Supporting the participation of English language learners in mathematical discussions. *For the Learning of Mathematics, 19*(1), 11-19.

Moschkovich, J. (2002). A situated and sociocultural perspective on bilingual mathematics learners. *Mathematical Thinking and Learning, 4*(2-3), 189-212.

Moschkovich, J. (2012, April). *Mathematics, the Common Core, and language: Recommendations for mathematics instruction for ELs aligned with the Common Core.* Paper presented at the Understanding Language Conference, Stanford, CA. http://ell.stanford.edu/papers

Moschkovich, J. (2015). Academic literacy in mathematics for English learners. *Journal of Mathematical Behavior, 40*, 43-62.

National Academies of Sciences, Engineering, and Medicine. (2018). *English learners in STEM subjects. Transforming classrooms, schools, and lives*. National Academies Press.

National Center for Education Statistics. (2019). *The condition of education 2019* (NCES 2019-144). U.S. Department of Education.

National Research Council. (1996). *National science education standards.* National Academies Press.

National Research Council. (2001). *Knowing what students know: The science and design of education assessment*. National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* National Academies Press.

National Research Council. (2014). *Developing assessments for the Next Generation Science Standards.* National Academies Press.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. N. Giere (Ed.), *Cognitive models of science: Minnesota studies in the philosophy of science*, Vol XV. University of Minnesota Press.

New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review, 66*, 60-92.

Next Generation Science Standards Lead States. (2013). *Next Generation Science Standards: For states, by states.* The National Academies Press.

O'Connor, M., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology and Education Quarterly, 24*(4), 318-335.

O'Halloran, K. (2008). *Mathematical discourse: Language, symbolism and visual images*. Continuum.

Otheguy, R., García, O., & Reid, W. (2019). A translanguaging view of the linguistic system of bilinguals. *Applied Linguistics Review, 10*(4), 625-651.

Pacheco, M., & Smith, B. (2015). Across languages, modes, and identities: Bilingual adolescents' multimodal codemeshing in the literacy classroom. *Bilingual Research Journal, 38*(3), 292-312.

Padalkar, S., & Ramadas, J. (2011). Designed and spontaneous gestures in elementary astronomy education. *International Journal of Science Education, 33*(12), 1703-1739.

Passmore, C. M., Gouvea, J. S., & Giere, R. (2014). Models in science and in learning science: Focusing scientific practice on sense-making. In M. R. Matthews (Ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1171-1202). Springer.

Passmore, C. M., Schwarz, C., & Mankowski, J. (2017). Developing and using models. In C. V. Schwarz, C. Passmore, & B. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 109-135). National Science Teachers' Association Press.

Patten, E. (2016). *The nation's Latino population is defined by its youth*. Pew Research Center.

Peirce, C. S. (1998). *The essential Peirce: Selected philosophical writings* (Vol. 3). Indiana University Press.

Penuel, W. R., & Watkins, D. A. (2019). Assessment to promote equity and epistemic justice: A use-case of a research-practice partnership in science education. *The Annals of the American Academy of Political and Social Science, 683,* 201-216.

Plummer, J. D., & Krajcik, J. (2010). Building a learning progression for celestial motion: Elementary levels from an earth-based perspective. *Journal of Research on Science Teaching, 47*(7), 768-787.

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 233-265.

Poehner, M. E., & van Compernoll, R. (2019, March). *Elaborating L2 dynamic assessment through praxis*. Paper presented at the Language Testing Research Colloquium, Atlanta, GA.

Prior, P. (2005). Moving multimodality beyond the binaries: A response to Gunther Kress' "Gains and Losses." *Computers and Composition, 22,* 23-30.

Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching, 49*(3), 363-393

Quinn, H., Lee, O., & Valdés, G. (2012, April). *Language demands and opportunities in relation to Next Generation Science Standards for English language learners: What teachers need to know*. Paper presented at the Understanding Language Conference, Stanford, CA. http://ell.stanford.edu/papers

Razfar, A., Khisty, L., & Chval, K. (2011). Re-mediating second language acquisition: A sociocultural perspective for language development. *Mind, Culture, and Activity, 18*(3), 195-215.

Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation, 37*, 15-24.

Ruiz-Primo, M. A., & Furtak, E. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment, 11*(3-4), 205-235.

Ruiz-Primo, M. A., & Furtak, E. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57-84.

Sasaki, M. (2000). Effects of cultural schemata on students' test-taking processes for cloze tests: A multiple data source approach. *Language Testing, 17*(1), 85-114.

Saunders, W. M., & Marcelletti, D. J. (2013). The gap that can't go away: The catch-22 of reclassification in monitoring the progress of English language learners. *Educational Evaluation and Policy Analysis, 35*(2), 139-156.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics approach*. Erlbaum.

Schwarz, C., & White, B. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction, 23,* 165-205.

Schwarz, C., Reiser, B., Acher, A., Kenyon, L., & Fortus, D. (2012). MoDeLS: Challenges in defining a learning progression for scientific modeling. In A. C. Alonzo & A. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 101-137). Sense.

Schwarz, C., Reiser, B., Davis, E., Kenyon, L., Acher, A., Fortus, D., . . . Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46*(6), 632-654.

Seah, L. H., Clarke, D. J., & Hart, C. E. (2011). Understanding students' language use about expansion through analyzing their lexicogrammatical resources. *Science Education, 95*(5), 852-876.

Shipka, J. (2013). Including, but not limited to, the digital: Composing multimodal texts. In T. Bowen & C. Whithaus (Eds.), *Multimodal literacies and emerging genres* (pp. 73-89). University of Pittsburgh Press.

Shohamy, E. (2011). Assessing multilingual competences: Adopting construct valid assessment policies. *The Modern Language Journal, 95*(3), 418-429.

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*(6), 864-881.

Siegel, M. A., Menon, D., Sinha, S, Promyod, N., Wissehr, C., & Halverson, K. L. (2014). Equitable written assessments for English language learners: How scaffolding helps. *Journal of Science Teacher Education, 25*(6), 681-708.

Smagorinsky, P. (2001). Rethinking protocol analysis from a cultural perspective. *Annual Review of Applied linguistics, 21,* 233-245.

Smith, B. (2019). Mediational modalities: Adolescents collaboratively interpreting literature through digital multimodal composing. *Research in the Teaching of English, 53*(3), 197-222.

Smith, B., Pacheco, M., & de Almeida, C. R. (2017). Multimodal codemeshing: Bilingual adolescents' processes composing across modes and languages. *Journal of Second Language Writing, 36,* 6-22.

Smith, J. P., diSessa, A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences, 3,* 115-163.

Smith, P. S., & Plumley, C. L. (2016). *A review of research literature on teaching about the small particle model of matter to elementary students*. Horizon Research.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge University Press.

Sobul, D. (1995). *Specially designed academic instruction in English.* ERIC (ED 391357).

Solano-Flores, G., Barnett-Clarke, C., & Kachchaf, R. R. (2013). Semiotic structure and meaning making: The performance of English language learners on mathematics tests. *Educational Assessment, 18*(3), 147-161.

Solano-Flores, G., Wang, C., & Shade, C. (2016). International semiotics: Item difficulty and the complexity of science item illustrations in the PISA-2009 international test comparison. *International Journal of Testing, 16*(3), 205-219.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Sage.

Tang, K. S., Delgado, C., & Moje, E. B. (2014). An integrative framework for the analysis of multiple and multimodal representations for meaning-making in science education. *Science Education, 98*(2), 305-326.

Tang, K. S., Tan, S. C., & Yeo, J. (2011). Students' multimodal construction of the work-energy concept. *International Journal of Science Education, 33*(13), 1775-1804.

Thurlow, M. L., & Kopriva, R. J. (2015). Advancing accessibility and accommodations in content assessments for students with disabilities and English learners. *Review of Research in Education, 39*, 331-369.

Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal, 27*(5), 615-631.

Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education, 24*(4), 357-368.

Turkan, S., Lopez, A., Lawless, R., & Tolentino, F. (2019). Using pictorial glossaries as an accommodation for English learner: An exploratory study. *Educational Assessment, 24*(3), 235-265.

Tytler, R., Prain, V., Hubber, P., & Waldrip, B. (2013). *Constructing representations to learn in science*. Sense.

Umansky, I. M., Thompson, K. D., & Díaz, G. (2017). Using an ever-English learner framework to examine disproportionality in special education. *Exceptional Children, 84*(1), 76-96.

Ünsal, Z., Jakobson, B., Wickman, P., & Molander, B. (2018). Gesticulating science: Emergent bilingual students' use of gestures. *Journal of Research in Science Teaching, 55*(1), 121-144.

Unsworth, L. (2006). Towards a metalanguage for multiliteracies education: Describing the meaning-making resources of language-image interaction. *English Teaching: Practice and Critique, 5*, 55-76.

U.S. Department of Education. (2012). *Elementary and Secondary Education Act Flexibility*. http://www2.ed.gov/polcy/elsec/guid/esea-flexibility/index.html

U.S. Department of Education. (2015). *Every Student Succeeds Act*. https://www.gpo.gov/fdsys/pkg/BILLS-114s1177enr/pdf/BILLS-114s1177enr.pdf

Valdés, G. (2004). Between support and marginalisation: The development of academic language in linguistic minority children. *International Journal of Bilingual Education and Bilingualism, 7*(2-3), 102-132.

Valdés, G. (2017). Entry visa denied: The construction of symbolic language borders in educational settings. In O. García, N. Flores, & M. Spotti (Eds.), *The Oxford handbook of language and society* (pp. 321-348). Oxford University Press.

van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective.* Kluwer Academic.

Vygotsky, L. S. (1998). The problem of age (M. J. Hall, trans.). In R. W. Rieber (Ed.), *The collected works of L. S. Vygotsky: Child psychology* (Vol. 5, pp. 187-205). Plenum.

Walqui, A., Koelsch, N., & Schmida, M. (2012). *Persuasion across time and space: Analyzing and producing complex texts.* http://ell.stanford.edu/teaching_resources

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology*, *25*(1), 127-147.

WIDA Consortium. (2012). *Amplification of the English language development standards*. Madison, WI: Board of Regents of the University of Wisconsin System. http://www.wida.us/standards/eld.aspx

WIDA Consortium. (2019). *Interpretive guide for score reports: Kindergarten–Grade 12*. Board of Regents of the University of Wisconsin System. https://wida.wisc.edu/sites/default/files/resource/Interpretive-Guide.pdf

Wilensky, U., Brady, C., & Horn, M. (2014). Fostering computational literacy in science classrooms. *Communications of the ACM*, *57*(8), 17-21.

Windschitl, M., & Thompson, J. (2006). Transcending simple forms of school science investigation: The impact of preservice instruction on teachers' understandings of model-based inquiry. *American Educational Research Journal*, *43*(4), 783-835.

Wolf, M. K., Guzman-Orth, D., Lopez, A., Castellano, K., Himelfarb, I., & Tsutagawa, F. (2016). Integrating scaffolding strategies into technology-enhanced assessments of English learners: Task types and measurement models. *Educational Assessment*, *21*(3), 157-175.

Yi, Y., King, N., & Safriani, A. (2017). Reconceptualizing assessment for digital multimodal literacy. *TESOL Journal*, *8*(4), 878-885.

Zhang, Y. (2016). Multimodal teacher input and science learning in a middle school sheltered classroom. *Journal of Research in Science Teaching*, *53*(1), 7-30.

Zuengler, J., & Miller, E. (2006). Cognitive and sociocultural perspectives: Two parallel SLA worlds? *TESOL Quarterly*, *40*(1), 35-58.