New content standards in English language arts and literacy, mathematics, and science expect all students, including English learners (ELs), to use multiple modes beyond language. Traditionally, conceptualizations of multimodality in EL education and the content areas have differed considerably. To acknowledge and begin confronting these differences, the author proposes weak and strong versions of multimodality. Whereas the weak version privileges language and takes nonlinguistic modes as scaffolds for language development with ELs, the strong version views multiple modes as essential to engaging in disciplinary practices. Adopting the strong version, the author analyzes work samples from a fourth-grade science classroom to illustrate how students deploy multiple modes strategically as they engage in scientific modeling. This analysis also shows how a multimodal lens can shed light on the unique affordances and limitations of language, a mode of particular relevance to ELs. The author argues that embracing the strong version of multimodality is not only necessary but transformative for ELs in the new content standards era, because it allows them to draw from the meaning-making resources at their disposal while engaging in disciplinary practices. In light of divergent conceptualizations of multimodality, the article ends with a call to action for closer collaboration between language and content educators.

doi: 10.1002/tesq.443

As a fast-growing population of English learners (ELs) is expected to meet rigorous content standards for college and career readiness, the field of EL education faces both opportunities and challenges. Recently, much attention has been paid to the language learning opportunities and challenges of new standards in English language arts (ELA) and literacy, mathematics, and science (e.g., Bailey & Heritage, 2014). However, these discussions have underemphasized the multimodal nature of new content standards, which call for
students to use multiple modes beyond language as they engage in disciplinary practices, such as arguing from evidence and analyzing texts. With new content standards expanding what “counts” as meaning-making in the content area classroom, there is a need to consider the implications of this shift for all students and for ELs in particular.

In recent years, theory and research in multimodality has gained increasing prominence. Bezemer and Kress (2008) define mode as “a socially and culturally shaped resource for making meaning” (p. 171). Examples of modes include, but are not limited to, images, videos, symbols, charts, tables, graphs, and oral/written language. Modality, then, refers to “the use of several semiotic modes in the design of a semiotic product or event” (Kress & van Leeuwen, 2001, p. 20).

The purpose of this article is to initiate a dialogue between the field of EL education and the content areas that acknowledges and confronts divergent conceptualizations of multimodality and their consequences for ELs. Historically, the field of EL education and the content areas have conceptualized multimodality differently. Until recently, these conceptualizations have coexisted in relatively unproblematic fashion. However, as new content standards expect all students, including ELs, to engage in disciplinary practices that are inherently multimodal, the fields are forced to confront these differences head on. Grounded in standards for reporting on humanities-oriented research in education (American Educational Research Association [AERA], 2009), I aim to “point out tensions and contradictions” between these conceptualizations that may “foster dissonance and discomfort with conventional practices” (AERA, 2009, p. 482) and, ultimately, serve as a catalyst for closer and more fruitful collaboration between EL education and the content areas. By inviting the fields into conversation around multimodality, this article takes an important step toward ensuring language and content educators are coordinated in their efforts to support ELs in the new content standards era.

I begin by describing traditional conceptualizations of multimodality in the field of EL education and the content areas, respectively. Based on these conceptualizations, I propose weak and strong versions of multimodality, arguing that the strong version is not only necessary but transformative for ELs in the new content standards era. Second, I describe how multimodality has been operationalized in new content standards and in current and emerging lines of work in EL education. Third, I analyze student work samples from a fourth-grade science classroom to show how students use multiple modes strategically to engage in a disciplinary practice. Finally, in light of these divergent conceptualizations of multimodality, I end with a call to action for closer and more substantive collaboration between EL education and content area education.
ELs represent the fastest-growing student population in the United States (Gándara & Hopkins, 2010). According to the 2010 U.S. Census (U.S. Census Bureau, 2012), 21% of children 5–17 years old spoke a language other than English at home. In 2013–2014, ELs constituted 9.3% of public school students, or an estimated 4.5 million students (National Center for Education Statistics, 2016).

In the midst of these demographic changes, a new wave of standards-based reform has raised the bar for learning in the content areas with the goal of preparing all students for college and careers. The Common Core State Standards (CCSS) for ELA/literacy and mathematics are implemented across most U.S. states. With regard to science standards, as of spring 2017, 18 states and the District of Columbia have adopted the Next Generation Science Standards (NGSS), and 14 additional states have adapted the NGSS for use as their state science standards. Together, these states represent more than 50% of the student population in the United States (Carnegie Corporation of New York, 2017). Even for the remaining states that may not adopt or adapt the NGSS, A Framework for K–12 Science Education (hereafter referred to as the Framework; National Research Council [NRC], 2012), which informed the development of the NGSS, is likely to influence science instruction. Although this article focuses specifically on new content standards in the United States, similar standards movements are taking place in other countries (see Kibler, Valdés, & Walqui, 2014, for a special issue of TESOL Quarterly dedicated to standards-based educational reform with ELs from an international perspective).

All three sets of standards previously mentioned promote an expanded view of meaning-making in the content areas by articulating disciplinary practices that are inherently multimodal. For example, students are expected to produce and interpret texts that incorporate multiple modes, such as graphs (science), formulas (mathematics), and multimedia (ELA). As a growing number of ELs are expected to engage in such multimodal disciplinary practices, divergent conceptualizations of multimodality between EL education and the content areas come into view. As a point of departure for understanding these differences, I begin by comparing terminology adopted by each field.

In EL education, the term modes (or modalities) has been used to describe the channels through which language is transacted. Modes can refer to either receptive and productive language (e.g., Council of Chief State School Officers [CCSSO], 2014) or the specific skills
contained within these, namely, listening and reading (receptive) and speaking and writing (productive; e.g., WIDA Consortium, 2012). In the content areas, the term *modes* (or *modes of representation*) has been used in a broader sense to describe the range of meaning-making devices used in the disciplines, both nonlinguistic and linguistic (NRC, 2012; New London Group, 1996). Examples from science include drawings, symbols, and graphs (nonlinguistic) as well as oral and written language (linguistic; NRC, 2012; Quinn, Lee, & Valdés, 2012). Thus, whereas EL education provides a definition primarily in terms of language, the content areas adopt a broader definition that includes, but is not limited to, language. This difference in terminology is worth highlighting, because it reflects the different foci and priorities of the fields that have informed their respective conceptualizations of multimodality. I refer to these as the *weak version of multimodality* and the *strong version of multimodality*, which are described in turn. Because the term *modes*, as defined by the content areas, is more inclusive, this is the definition I rely on throughout the article.

The Weak Version of Multimodality

Although multimodality is featured prominently in new content standards, it has not always been accorded this same status in EL education. With ELs, nonlinguistic modes are typically seen as supports for language development. They are a crutch or temporary scaffold to be removed once students develop proficiency with more privileged forms of communication, namely, oral and written language. Although this focus on language is intuitive given the field’s commitment to supporting ELs in learning English, it has also served to perpetuate a longstanding tradition of logocentrism that takes language to be “the standard of precision in meaning” and all other modes to be inherently ambiguous (Lemke, 2002, p. 31). Consider, for example, a recommendation from the Sheltered Instruction Observation Protocol (SIOP) model (Echevarría, Vogt, & Short, 2012), a widely adopted framework for addressing the academic needs of ELs in the content areas:

Allow students alternative forms for expressing their understanding of information and concepts. Often English learners have learned the lesson’s information but have difficulty expressing their understanding in English, either orally or in writing. Hands-on activities can be used to reinforce the concepts and information presented, with a reduced linguistic demand on these students. In a high school science class,
students demonstrated their understanding of concepts such as hydroplaning by drawing a sketch of a car on a wet road and labeling the drawing to show their knowledge.

(p. 102)

This recommendation illustrates the weak version of multimodality in several ways. The use of the term *alternative* positions language as central to science learning, whereas other modes (e.g., drawing) exist on the periphery. Nonlinguistic modes are meant to “reinforce” learning, but they are not the preferred modes of expression, offering no particular affordances beyond “reduced linguistic demand.” Thus, ELs’ sketches are valued primarily for their compensatory function and are considered to be no longer necessary once these students can demonstrate their understanding using language. The underlying assumption of this view is that nonlinguistic modes are “illustrative supports to the ‘real thing’” (Kress, Jewitt, Ogborn, & Tsatsarelis, 2014, p. 51) rather than legitimate meaning-making resources in their own right.

In a short commentary published by *TESOL Quarterly*, Kress (2000) challenged this assumption by encouraging EL educators to reconsider what “counts” as meaning-making in their classrooms. Specifically, he criticized EL educators for “act[ing] as though language fully represented the meanings they wish to encode and communicate” while viewing other modes as “someone else’s to look after” (p. 337). However, nearly two decades after Kress’s commentary, the weak version of multimodality continues to be the prevailing view in EL education (see “Multimodality in EL Education” section for specific examples).

**The Strong Version of Multimodality**

 Whereas the weak version of multimodality is the prevailing view in EL education, the strong version is more commonly adopted in the content areas. The strong version retains some assumptions of the weak version but also extends them to account for norms and conventions of the disciplines. It asserts that, although nonlinguistic modes can assist ELs in communicating their ideas, particularly at the early stages of language proficiency, this represents a narrow and reductive conception of multimodality. Modes are not simply scaffolds or supports; they are the essential semiotic tools of the disciplines. In science, for example, graphs, tables, and charts are central to communicating scientific meanings (Lemke, 1998). Each of these modes has particular affordances and limitations, or potentials and constraints, for making meaning (Bezemem & Kress, 2008). In the example above, the sketch of the car may assist ELs in communicating their
understanding of hydroplaning, but it may also afford opportunities for all students to represent the concept more effectively and to engage in the type of visual representation that is highly valued in science (Mathewson, 1999).

Although multimodality has long been recognized in the content areas, its importance is magnified by new content standards that are inherently multimodal. As students engage in disciplinary practices of content standards, they make strategic and deliberate use of multiple modes, taking account of the affordances and limitations of each mode as well as norms and conventions of each discipline (see “Multimodality in New Content Standards” section for specific examples).

A Case for the Strong Version of Multimodality in the New Content Standards Era

Although the strong version is deeply ingrained in the content areas, it has not been widely embraced in EL education, where the conventional notion of nonlinguistic modes as scaffolds for learning language (i.e., the weak version) has overshadowed their disciplinary significance. Such a persistent focus on language (especially “academic language”) has often resulted in the exclusion of ELs from meaningful participation in rigorous content area instruction (Valdés, 2017). However, in the new content standards era, strategic use of multiple modes in ways appropriate to each discipline is itself a goal of learning. With this shift, what was once compensatory for ELs is now expected of all students, regardless of proficiency in English.

Embracing the strong version of multimodality has far-reaching consequences, because it turns a deficit view of ELs as lacking the linguistic resources needed to participate meaningfully in the content areas into an asset view that capitalizes on the diverse meaning-making resources they bring to the classroom. Thus, in the new content standards era, the strong version of multimodality is not only necessary but transformative for ELs, because it allows them to draw from their full semiotic repertoire while engaging in disciplinary practices called for by new content standards.

At this point, it is important to emphasize that, for ELs, developing proficiency in English is vital to their school success and an important goal of their education. Thus, the strong version of multimodality does not imply a disregard for or devaluing of language. Much the opposite, adopting a multimodal lens, as I propose in this article, is largely motivated by the integral role of language in communication. However, in response to the monomodal focus of much
previous work, I suggest that broadening our gaze beyond language can help elucidate the distinct semiotic potential of language in relation to other modes. In other words, if we “step outside [of language] and take a satellite view of it” (Kress et al., 2014, p. 10), we can sharpen our focus on the unique contribution of language to meaning-making in the content areas. In a later section, I show how adopting a multimodal lens can help shed light on the particular affordances and limitations of written language in student work samples from an NGSS science classroom.

In sum, I propose a weak and a strong version of multimodality in EL education and the content areas, respectively. Whereas the weak version privileges language and takes nonlinguistic modes as scaffolds for language development with ELs, the strong version views multiple modes as essential to engaging in disciplinary practices. Table 1 summarizes the key features of the weak and strong versions of multimodality in relation to three points of divergence: (1) Who uses multiple modes? (2) How are multiple modes used? (3) Which modes are valued? In light of new content standards that are inherently multimodal, I argue that adopting the strong version of multimodality is not only necessary but transformative for ELs.

MULTIMODALITY IN TEACHING AND LEARNING NEW CONTENT STANDARDS WITH ELS

In this section, I provide illustrative examples of how multimodality has been operationalized in new content standards and EL education. I call attention to the weak and strong versions of multimodality as they appear in these examples.

| TABLE 1 |
| Key Features of the Weak and Strong Versions of Multimodality |
| Weak version of multimodality | Strong version of multimodality |
| Who uses multiple modes? | ELs use multiple modes until they have developed proficiency in English. | All students, regardless of EL status or proficiency in English, use multiple modes to engage in disciplinary practices. |
| How are multiple modes used? | Use of multiple modes is seen as compensatory. | Use of multiple modes is seen as strategic. |
| Which modes are valued? | Language is the default privileged mode. | All modes are valued based on their particular affordances and limitations as well as norms and conventions of each discipline. |

MULTIMODALITY AND ENGLISH LEARNERS
Multimodality in New Content Standards

Recent scholarship on multimodality has provided the conceptual foundation and associated terminology for developing multimodal content standards. This work has been informed by sociosemiotic theories of communication that take language and other meaning-making modes to be a system of choices (Halliday, 1978). Specifically, research on multimodality has emphasized the unique affordances and limitations of different modes and how these can be leveraged by students and teachers to “mean more, mean new kinds of meanings never before meant and not otherwise mean-able” (Lemke, 1998, p. 93). From this perspective, an image can never do the same semiotic work as a written text, and vice versa. When used in combination, image and text join forces to create greater meaning potential than would be possible with either mode alone. In this way, multimodal communication is the “logical product, in a multiplicative sense, of the capacities of the constituent semiotic resource systems” (Lemke, 2002, p. 303). Because multimodal communication requires strategic choices that take account of the affordances and limitations of each mode, there is an increasing trend toward characterizing this process as design, or “the intentional deployment of resources in specific configurations to implement the purposes of the designers” (Kress, 2000, p. 340). Although expanding what “counts” as meaning-making is important for all students, it may be especially beneficial for students from diverse backgrounds, whose cultural and linguistic resources are recruited as valued resources for learning (New London Group, 1996; see Smith, 2014, for a review of research on multimodal writing with diverse student groups).

Key concepts from research on multimodality (e.g., affordances, limitations, design) figure prominently in new content standards. Whereas previous content standards focused on students’ mastery of discrete concepts and skills, a distinctive feature of new content standards is their focus on disciplinary practices, or practices “in which students and teachers engage to construct knowledge, concepts, and skills in particular subject areas” (CCSSO, 2012, p. 1). Engaging in disciplinary practices called for by the CCSS and NGSS requires that all students make strategic use of multiple modes in ways appropriate to each discipline. In what follows, I present illustrative (but not exhaustive) examples of multimodal disciplinary practices from the CCSS for ELA/literacy, CCSS for mathematics, and NGSS. For each set of standards, a different feature of the strong version of multimodality is foregrounded (see Table 1). However, because the features are interrelated, they are present in all three sets of standards.
Who uses multiple modes in the NGSS?. The eight science and engineering practices articulated by the NGSS are inherently multimodal and reflect how scientists and engineers carry out their work (NGSS Lead States, 2013). In the NGSS science classroom, all students are expected to develop and use models, analyze and interpret data, use mathematics and computational thinking, and obtain, evaluate, and communicate information. As students engage in these practices, they draw on multiple modes, including symbols, equations, tables, charts, and graphs. They also use multiple modes to engage in practices traditionally considered language based, such as arguing from evidence and constructing explanations. For example, students may develop model-based arguments or explanations about science phenomena.

Although all students use multiple modes to engage in NGSS science and engineering practices, ELs in particular may benefit from the multimodal nature of these practices. As part of its vision for inclusive science instruction, the Framework advocates that “opportunities [for ELs] to edit or to display their knowledge in less language-embedded tasks would help level the playing field” (NRC, 2012, p. 289). Thus, the Framework and NGSS view multimodality as both indispensable to engagement in science and engineering practices and particularly beneficial to ELs.

How are multiple modes used in the CCSS for ELA/literacy?. Spurred by the rapid proliferation of information and multimedia technologies, the CCSS for ELA/literacy (National Governors Association Center for Best Practices [NGA] & CCSSO, 2010a) posit an expanded view of language that positions students as strategic designers and interpreters of multimodal meanings (Cope & Kalantzis, 2009; New London Group, 1996; Street, 2003). In the standards for speaking and listening, for example, students across all grade levels are expected to “make strategic use of digital media (e.g., textual, graphical, audio, visual, and interactive elements) in presentations” (NGA & CCSSO, 2010a, p. 22). Similarly, in writing, students are expected to strategically deploy multiple modes, including graphics, formatting, multimedia, and written language, to meet “varying demands of audience, task, and purpose” (NGA & CCSSO, 2010a, p. 7). Furthermore, in reading, students are expected to examine the strategic choices of other writers, for example, by “comparing and contrasting a text to an audio, video, or multimedia version of the text, analyzing each medium’s portrayal of the subject” (NGA & CCSSO, 2010a, p. 39). By characterizing students’ use and interpretation of multiple modes as strategic, the CCSS for ELA/literacy elevate
the status of multimodality to a valued learning outcome beyond a scaffold.

**Which modes are valued in the CCSS for mathematics?** The CCSS for mathematics articulate eight mathematical practices that are inherently multimodal (NGA & CCSSO, 2010b). For example, one practice expects that students use appropriate tools strategically, where tools include modes such as tables, graphs, flowcharts, and formulas. To carry out this practice, students must consider the affordances and limitations of different modes in deciding when and how to use them:

Mathematically proficient students . . . are sufficiently familiar with tools appropriate for their grade or course to make sound decisions about when each of these tools might be helpful, recognizing both the insight to be gained and their limitations.

(NGA & CCSSO, 2010b, p. 7)

In conjunction with considering affordances and limitations, students must also become proficient with conventional representations of mathematical ideas, such as coordinate planes and algebraic notation (NGA & CCSSO, 2010b). Such disciplinary ways of making meaning are critical to engaging in mathematical practices, such as modeling with mathematics, even those practices traditionally considered language based, such as constructing viable arguments.

In sum, the NGSS and CCSS for ELA/literacy and mathematics embrace the strong version of multimodality by expecting all students, including ELs (i.e., who uses multiple modes?), to use multiple modes strategically (i.e., how are multiple modes used?), taking account of the affordances and limitations of each mode as well as norms and conventions of each discipline (i.e., which modes are valued?; see Table 1). By articulating multimodal disciplinary practices as important goals of learning, all three sets of standards send a clear message affirming the value of multimodality in their respective disciplines.

**Multimodality in EL Education**

In this section, I address multimodality in EL education. First, drawing on examples from the set of English language proficiency (ELP) standards adopted by the majority of states in the nation, I show how the weak version of multimodality continues to dominate EL education. Next, I offer examples of emerging lines of work that suggest a shift toward the strong version of multimodality, including a more recent set of ELP standards.
Current work representing the weak version of multimodality. By and large, EL education has adhered to the weak version of multimodality, taking nonlinguistic modes to be scaffolds for language development with ELs. Adopted by 35 states and four districts, territories, and agencies across the United States, the WIDA English Language Development (ELD) Standards address the language needed to “communicate information, ideas and concepts necessary for academic success” (WIDA Consortium, 2012, p. 3) in each of four content areas: language arts, mathematics, science, and social studies. The bulk of the ELD Standards revolves around the word/phrase-, sentence-, and discourse-level features of academic language expected of students at each level of English proficiency, and nonlinguistic modes are described as instructional supports for scaffolding language development. The ELD Standards provide a menu of modes in the form of sensory supports (e.g., drawings, models, photographs) and graphic supports (e.g., charts, tables, timelines) for teachers to draw on flexibly when planning curriculum, instruction, and assessment for ELs (WIDA Consortium, 2012, p. 11). The assumption underlying this conceptualization is that “communicating information, ideas, and concepts” (WIDA Consortium, 2012, p. 3) in the content areas can be done with language alone and that, once ELs develop proficiency with language, they no longer require the use of multiple modes. Although this focus on language is to some degree expected given WIDA’s charge (i.e., to develop English proficiency standards), it raises questions about whether the WIDA ELD Standards support ELs to engage in multimodal disciplinary practices of content standards.

As an example of how WIDA operationalizes multimodality, Figure 1 shows a model performance indicator (MPI) from the WIDA ELD Standards. MPIs are examples of how teachers might scaffold instruction and assessment for ELs at each proficiency level. This sixth-grade MPI addresses an NGSS standard related to the functions of organisms within ecosystems. As shown in the figure, students at Level 1 proficiency are expected to “identify functions of organisms within ecosystems based on oral statements using photos or illustrations” (WIDA Consortium, 2012, p. 85, emphasis added). According to the progression articulated in the MPI, by the time students reach Level 3 proficiency and beyond, they no longer require the use of photos and illustrations. Students at Level 5 proficiency, for example, are expected to categorize the functions of organisms using language alone (“based on extended oral discourse”). Implicit in this progression is that the visual mode (e.g., illustrations) is not central to science learning, but a means to an end for scaffolding language development. Meanwhile, the NGSS standard referenced in the MPI (see “Connection” in the figure) requires that students “develop a model to describe the cycling
of matter and flow of energy among living and nonliving parts of an ecosystem” (MS-LS2-3). Because engaging in the NGSS practice of modeling requires multiple modes, regardless of a student’s English proficiency level, the MPI does not accurately reflect the multimodal nature of this science practice.

Emerging lines of work moving toward the strong version. Despite the predominance of the weak version of multimodality, there is some indication that EL education is moving toward the strong version. This shift has occurred in the wake of the Every Student Succeeds Act (ESSA) of 2015, which mandates that each state “adopt English language proficiency standards that... are aligned with challenging state academic standards” (U.S. Department of Education, 2015, p. 24, emphasis added). The implication of this federal legislation is that ELP standards must reflect the language underlying successful mastery of content standards and, more specifically, the language needed to engage in disciplinary practices. Because engaging in disciplinary practices involves the use of nonlinguistic modes in conjunction with language, ELP standards must adapt to reflect this change.

A shift toward the strong version of multimodality was spearheaded by the work of the Understanding Language Initiative at Stanford University (http://ell.stanford.edu), devoted to improving education for ELs in light of the CCSS and NGSS. As part of this initiative, scholars and policymakers from both EL education and the content areas collaborated to produce a series of commissioned papers that addressed the opportunities and demands presented by the CCSS for
ELA/literacy (Hull & Moje, 2012), CCSS for mathematics (Moschkovich, 2012), and NGSS (Quinn et al., 2012). A common theme emerging from these papers is the critical role of multimodality in meeting new content standards. Furthermore, among several instructional shifts promoted by the initiative was a shift from “traditional grammar as a starting point for students to know about language” to “multimodal grammar... to support students’ understanding of the visual, spatial, gestural, audio and linguistic meanings of texts” (Walqui, Koelsch, & Schmida, 2012, p. 2).

Based on the work of the Understanding Language Initiative (Lee, Quinn, & VALDÉS, 2013), the CCSSO (2012) developed a framework for evaluating “correspondence” (now called “alignment” under ESSA) between ELP standards and the CCSS and NGSS to provide guidance to states. As one of its eight guiding principles, the framework recommends that “ELP standards explicitly cultivate students’ ability to comprehend and communicate about complex text in oral, written, visual, and multimodal forms” (CCSSO, 2012, p. 25). Thus, the CCSSO framework recognizes multimodality as on par with, rather than subservient to, oral and written language.

The influence of the CCSSO (2012) framework is evident in the ELP standards recently developed by the multistate consortium, English Language Proficiency Assessment for the 21st Century (ELPA21). Adopted by 10 states across the United States, the ELPA21 ELP Standards (CCSSO, 2014) assert that “relevant, strategic, and appropriate multimedia tools... aligned to the ELP Standards should be integrated into the design of curriculum, instruction, and assessment for ELLs” (p. 3). Standard 1, for example, calls for students to “construct meaning... through grade-appropriate listening, reading, and viewing” (p. 4, emphasis added). By including “viewing” alongside listening and reading, the Standards recognize the visual mode as a legitimate source of meaning in the content areas. However, whereas Standard 1 is inclusive of multimodality, Standards 3 and 4 are not, as they address speaking and writing in the content areas but not “representing.” As described above, modeling is a valued form of representation for communicating ideas in science through the use of multiple modes, including drawings and symbols as well as written language. Thus, the ELPA21 ELP standards show some signs of the strong version of multimodality while also retaining aspects of the weak version.

The increasing visibility of sociocultural theories of learning has also expanded traditional conceptions of meaning-making. As one example, vanLIER’s (2002) ecological-semiotic perspective challenges narrow views of nonlinguistic modes as “trivial add-ons,” arguing that “it may make as much sense to separate language from other semiotic processes as it does to separate the swaying of the tree branch from the
wind that moves it” (p. 149). Applying this perspective in the context of new content standards, vanLier and Walqui (2012) propose a view of language as action that is “intimately connected to all other forms of action, physical, social and symbolic” (p. 47). From an action-based perspective, students put all of their meaning-making resources in the service of “doing” the work of the content areas. Thus, rather than seeing language as an autonomous system, sociocultural theories view language as fully integrated with and inseparable from other modes.

Multimodality has also gained increasing prominence in bilingual education, most notably in García and colleagues’ notion of translanguaging (e.g., García & Li, 2014). This work is premised on the idea that the boundaries between meaning-making systems are social constructions that do not reflect the fluid and dynamic nature of students’ actual communicative practices. Translanguaging extends beyond language to include all meaning-making modes, as students select features strategically from their full linguistic and semiotic repertoire to communicate effectively. Because no single mode is privileged over another and “all semiotic practices are licensed as valuable” (Otheguy, García, & Reid, 2015, p. 302), students use a range of “multilingual multimodal practices” (García & Li, 2014, p. 134) in flexible combinations to access and engage with rigorous content. This work has not explicitly addressed the role of multimodality in the content areas or in new content standards (see Smith, Pacheco, & Rossato de Almeida, 2017, for a recent exception in ELA), but it offers valuable insight into the potential for learning when all students, and ELs in particular, are supported to utilize their full semiotic repertoire.

Most recently, a special issue of TESOL Quarterly titled “Multimodality: Out From the Margins of English Language Teaching” invited EL educators to critically consider the role of multimodality in language teaching and learning (Early, Kendrick, & Potts, 2015). This special issue covered a range of topics, from multimodal identity texts (Cummins, Hu, Markus, & Montero, 2015) to digital literacy practices (Toohey et al., 2015), with the goal of “bring[ing] issues of multimodality and meaning making squarely at the center of TESOL’s concerns” (Early et al., 2015, p. 451).

In sum, current work in EL education, represented by the WIDA ELD Standards, exhibits key features of the weak version of multimodality by expecting ELs to use multiple modes in a compensatory manner (i.e., how are multiple modes used?) until they have developed proficiency in English (i.e., who uses multiple modes?), which is the default privileged mode (i.e., which modes are valued?; see Table 1). However, emerging lines of work in EL education suggest a shift toward the strong version. Building on this momentum, while also adding a disciplinary perspective, I advocate a shift toward the
strong version of multimodality that posits language as one mode, among many, needed to engage in rigorous learning in the content areas.

AN EXAMPLE OF MULTIMODALITY FROM THE NGSS SCIENCE CLASSROOM

Although multiple modes are found in all content areas, they are especially pervasive in science (e.g., Lemke, 1998). For this reason, the NGSS science classroom offers a rich window into the transformative potential of adopting the strong version of multimodality, particularly for ELs. Using student work samples from an NGSS science classroom, I show how modes are much more than supports for ELs; they are essential semiotic resources for communicating scientific meaning. The main purpose of this section is to highlight the strategic choices made by students as they engage in a multimodal disciplinary practice, but the analysis also offers insight into the affordances and limitations of one particular mode: language. By stepping away from language and adopting a multimodal perspective, we are able to see more clearly the semiotic work that language accomplishes in the science classroom.

To provide some context for the student work samples, an NGSS performance expectation in physical science for fourth grade is presented in Figure 2. This performance expectation blends a science and engineering practice (“develop a model”), a disciplinary core idea (“that light reflecting from objects and entering the eye”), and a crosscutting concept (cause and effect; “light... allows objects to be seen”). The blending of the three dimensions represents NGSS three-dimensional learning (NGSS Lead States, 2013).

Students were given a box with two holes (an eye hole and a light hole covered by a flap). Inside the box was a toy turtle affectionately named Tuck by students in the class. Peering into the box, students
made observations of the turtle under four conditions: (1) flap closed, (2) flap open, (3) flashlight shining through flap, and (4) lid of box open. When they were unable to see the turtle with the flap closed (condition 1) but were able to see it with the flap open (condition 2) and even more clearly with the flashlight shining through the flap (condition 3) and the box open (condition 4), they came to understand that light causes an object to be seen (crosscutting concept) when it reflects off the object and enters the eye (disciplinary core idea). Students were asked to develop models (science and engineering practice) of how they could see the toy turtle in the box.

Developing and using models is a key practice in the NGSS. According to the Framework, models are “representations that are in some ways analogous to the phenomena they represent” (NRC, 2012, p. 56) and can include physical replicas, diagrams, mathematical representations, and computer simulations. A key shift of the NGSS is that models show not only the components of a system but also the relationships among those components, thus enabling students to develop causal explanations. In their lightbox models, students are expected to represent the light source (flashlight), object (turtle), and eye (i.e., the components) as well as the process by which light originates from the flashlight, reflects off the turtle, and enters the eye (i.e., the causal relationship among the components). Modeling in the NGSS stands in contrast with traditional approaches to modeling in science class where students create one-off drawings with labeled components (e.g., a cell with each cell part labeled).

Because all models are abstractions that “bring certain features into focus while minimizing or obscuring others” (NRC, 2012, p. 56), modeling requires strategic decisions about what components and relationships to include or exclude, given the model’s purpose, and how to represent those components and relationships effectively, based on the affordances and limitations of each mode as well as disciplinary norms and conventions. Such decisions and trade-offs in design make modeling especially fertile ground for examining the strategic use of multiple modes in the NGSS science classroom.

**Lightbox Models**

In this section, I analyze the lightbox models of two students from the perspective of the strong version of multimodality. In the first model, I highlight the strategic use of multiple modes, including drawing, symbols, and language, in ways recognized by the discipline. In the second model, I devote more focused attention to the affordances and limitations of written language, a mode of particular relevance to
ELs. In this model, written language affords the potential for conveying sequential relations of events, but it does not convey the spatial relations among components, which is instead communicated through drawing. Such affordances and limitations have important consequences for how students represent different aspects of their thinking while engaging in the practice of modeling.

**Student A’s model.** Figure 3 shows Student A’s lightbox model. The student represents the lightbox from a side view to display what is happening inside. The model includes key components from the investigation, including the flashlight, turtle, and eye, and the black coloring of the lid helps identify the condition of the investigation represented in the model (i.e., flashlight with lid closed). The eye is represented not as a part of a face or body but as an isolated component, signaling a strategic choice to focus on one particular component (i.e., the eye) of a much larger system (i.e., the human body). The fact that light is drawn as straight lines (rather than curves) suggests an understanding of how light travels, and the use of dashed (rather than solid) lines suggests an understanding of light as invisible.

Moreover, the model uses arrows as symbols to show the causal relationship among the components. Specifically, it shows the path of light originating from the flashlight, passing through the open flap, reflecting off the turtle, and entering the eye. The science idea of reflection is communicated by the direction of the arrow, which is pointing toward the eye after reflecting off the turtle. (A common misconception is to draw an arrow in the opposite direction from the eye pointing toward the turtle.) In science, arrows are key semiotic resources for conveying conceptual relations, especially causality (Kress et al., 2014).

In addition to the drawing and symbols, the model includes written language in the form of labels and a title. Specifically, this student labels the light source (“Light”) and the path of light from the turtle to the eye using disciplinary terminology (“Reflection”). Although drawing and symbols carry the bulk of meaning in this model, written language is used strategically to highlight the most essential components and relationships. Overall, Student A demonstrates a well-developed understanding of reflection and the causal relationship using a strategic combination of multiple modes in ways recognized by the discipline.

**Student B’s model.** As shown in Figure 4, Student B uses drawing and symbols in a similar manner, showing the path of light originating from the flashlight, passing through the flap, reflecting off the turtle,
and entering the eye. In this model, however, written language figures more prominently, as the student provides a written account of what is happening in the lightbox at the top of the model (“Light travels in through the flap and reflects off Tuck into our eye so we can see Tuck”). The more extensive use of written language in this model raises the question of what distinct semiotic work language accomplishes beyond what is communicated with drawing and symbols. In other words, what are the affordances of written language in this model? The next logical question is whether drawing and symbols accomplish particular semiotic work that written language does not. In other words, what are the limitations of written language in this model? These affordances and limitations are discussed next.

**Affordances of written language.** In contrast to drawing, which is nonsequential, written language is founded on the “logic of succession in time” (Kress, 2000, p. 339), or the sequential ordering of events. As such, written language offers unique affordances for conveying sequential relations by specifying what happens first, second, third, and so on. Student B uses written language to describe the ability to see an object as resulting from three sequentially ordered events: “Light travels in through the flap” (event 1) and then “reflects off Tuck into our
eye” (event 2) and then “we can see Tuck” (event 3). The sequential ordering of events is especially important for conveying causality, as cause (i.e., light reflecting off an object) temporally precedes effect (i.e., the ability to see the object). This causal relationship is made explicit by the use of the coordinating conjunction so, which links the second and third events. When viewed in the context of its larger multimodal ensemble, written language can be seen as creating a “reading path” (Bezemer & Kress, 2008, p. 185) that engages the “reader” with the drawing and symbols in an order determined by the student. Without written language, the “reader” must resort to some degree of extrapolation to interpret the causal sequence of events represented by the drawing and symbols.

**Limitations of written language.** With written language offering unique affordances for conveying causal sequences of events, the question becomes whether drawing and symbols are no longer necessary. In other words, are the drawing and symbols merely replicating the semiotic work that language has already accomplished? From the perspective of the strong version of multimodality, the answer is unequivocally no, because language, as with any mode, has limitations. As one example, a limitation of written language in this model (and also in
the first model above) is that it does not communicate how the components of the lightbox (flashlight, turtle, eye) are positioned relative to each other. The spatial relations among components is important in this model, because students are expected to understand that in order for reflection to occur, there must be a clear and unobstructed path from the light source to the object and from the object to the eye. Because written language does not require a “commitment” (Bezemer & Kress, 2008, p. 176) to the spatial relations among components, it does not communicate the idea of an unobstructed path.

Drawing, on the other hand, requires a commitment to the spatial relations among components, thus enabling students to make visible an aspect of their thinking that would not be apparent based on language alone. This is because drawing, unlike language, is founded on “the logic of display in space” (Kress, 2000, p. 339). Whereas Student A communicates the idea of an unobstructed path by arranging the components in a triangular formation, Student B communicates this idea in an even more sophisticated manner by representing the components in a row and then shading the obstructed area where a shadow would likely result (see shading behind the turtle in Figure 4). Thus, conveying spatial relations is a limitation of written language but an affordance of drawing, and the two modes work in tandem to communicate more than could be communicated with any singular mode.

DISCUSSION

As a growing population of ELs is expected to meet new content standards that are inherently multimodal, divergent conceptualizations of multimodality in EL education and the content areas are brought into focus. To summarize, these divergent conceptualizations form the basis of what I refer to as the weak and strong versions of multimodality. In the weak version, nonlinguistic modes are scaffolds to be removed once ELs develop proficiency with oral and written language. In the strong version, multiple modes are the semiotic resources necessary for engaging in disciplinary practices and are especially beneficial to ELs. Whereas the weak version is a logical extension of EL educators’ commitment to promoting language development, the strong version recognizes that, when the goals of learning go beyond language development alone to engagement in multimodal disciplinary practices, there is a need to expand what “counts” as meaning-making in the content area classroom.

Using models from an NGSS science classroom, I show how two students deploy multiple modes strategically as they engage in the science disciplinary practice of developing models. Specifically, they make
purposeful decisions about what perspective to take (e.g., a side view), what components to include or exclude, and how to represent those components and their relationships effectively. In making these decisions, they consider the affordances and limitations of different modes as well as norms and conventions of science disciplines. For example, the use of arrows features prominently in both students’ models, as it allows them to communicate causality in a way that is consistent with disciplinary norms and conventions (Kress et al., 2014). Though modeling may be particularly beneficial to Student A, who is able to convey a great deal of meaning and complex thinking using mostly drawing and symbols, all students in the NGSS science classroom, regardless of English proficiency, engage in this multimodal disciplinary practice.

Furthermore, Student B’s lightbox model offers additional insight into the unique affordances and limitations of written language. As Kress et al. (2014) point out, “if language does not carry all the meaning in a communicative situation, it becomes crucial to know exactly which meanings it does” (p. 211). The sequential nature of written language allows Student B to communicate a sophisticated understanding of the causal relationship between reflection of the flashlight and the eye’s ability to see the turtle. Given the many affordances of written language, it is clear that developing proficiency in English must be considered an important goal of learning for all students in the content areas. This is a crucial and perhaps reassuring point for EL educators concerned that an emphasis on multimodality means diverting attention away from language development. The key difference from the weak version of multimodality is that the strong version values language not simply because it is the default privileged mode, but because of its unique affordances for communicating disciplinary meanings.

At the same time, language has limitations. Even in Student B’s model, which boasts extensive and sophisticated use of written language, the model conveys more with the combination of drawing, symbols, and written language than with any single mode. Written language is effective for conveying sequential relations but falls short of communicating spatial relations among components in the lightbox, which is communicated through drawing. This mutual complementarity creates synergy across modes and produces a multiplicative effect on the model’s meaning potential (Lemke, 2002). It is expected that even after ELs have attained proficiency in English, they continue to draw on multiple modes, because multimodality allows them to engage effectively in disciplinary practices. This expectation is at the heart of the strong version.

Building on momentum generated by new content standards as well as emerging lines of work in EL education, I argue that embracing the
strong version of multimodality is not only necessary but transformative for ELs, because it allows them to draw from their full semiotic repertoire while engaging in disciplinary practices. In the strong version, even beginner ELs, who use minimal language (as in Student A’s case), are positioned as strategic users of multiple modes and as fully capable of engaging in disciplinary practices called for by new content standards. Thus, rather than being compensatory for ELs, multimodality becomes validating and empowering. On the other end of the spectrum, even students who are proficient users of English (as in Student B’s case) use multiple modes because of the unique affordances of these modes for communicating disciplinary meanings.

IMPLICATIONS FOR PRACTICE, POLICY, AND RESEARCH

As new content standards take hold in schools and classrooms across the United States, both opportunities and challenges become apparent. To ensure that ELs capitalize on these opportunities and meet these challenges, there is an urgency at multiple levels of the education system for interrogating conventional practices (e.g., the weak version of multimodality) and envisioning new ones (e.g., the strong version of multimodality). Although this article is situated in the context of new content standards in the United States, the issues at stake are a matter of disciplinary, rather than national, boundaries and therefore have relevance beyond the U.S. context.

The strong version of multimodality offers important implications for practice, policy, and research. With regard to practice, the strong version urges teachers to see their students’ multimodal products as intentionally designed and imbued with meaning. As shown in the lightbox models, each decision made by students (e.g., color, line type, direction of arrows, position of components) provides a window into different aspects of their thinking. From a policy perspective, the strong version suggests a need for ELP consortia as well as individual states to consider how their standards and associated assessments align, as mandated by ESSA, with multimodal content standards. Furthermore, researchers may consider how examining language “as it is nestled and embedded within a wider semiotic frame” (Jewitt, 2017, p. 2) can help highlight the language ELs must acquire to communicative effectively in the new content standards era.

Enacting shifts in practice, policy, and research cannot be accomplished with educators working in silos. The divergent conceptualizations of multimodality outlined in this article are the product of a
much broader issue concerning a lack of mutual understanding and productive dialogue between EL education and the content areas. As a growing population of ELs is expected to achieve rigorous content standards, EL educators must become familiar with the ways multiple modes are used to make meaning in the content areas, just as content educators must become familiar with the ways multiple modes can be leveraged to support ELs in accessing content instruction in English. It is only through such sustained and substantive collaboration that multimodality can realize its full potential as indispensable to engagement in disciplinary practices and especially beneficial to ELs.

ACKNOWLEDGMENTS

I would like to acknowledge Rita Januszyk, who kindly shared her fourth-grade students’ work samples from the lightbox investigation task. I would also like to acknowledge that the lightbox investigation task is based on a similar task for middle school science in the Investigating and Questioning Our World Through Science and Technology curriculum units.

THE AUTHOR

Scott Grapin is a doctoral candidate in TESOL in the Department of Teaching and Learning at New York University’s Steinhardt School of Culture, Education, and Human Development. His research interests include the integration of content and language learning, second language writing, and language assessment.

REFERENCES


