Contemporary Classroom Vocabulary Assessment for Content Areas

Katherine A. Dougherty Stahl, Marco A. Bravo

The research-based content area vocabulary assessments presented in this article can help teachers feel confident in taking an assertive stance in developing vocabulary assessments based on their own curriculum needs.

Osa (all names are pseudonyms) teaches third grade in a high-poverty urban setting with a diverse population that includes a majority of children of color and a high percentage of English-language learners (ELLs). During the most recent school year, she instructed vocabulary in a deliberate way during the literacy block and content area instruction. In light of the increased time and attention to vocabulary instruction, she felt confident that her students had increased word knowledge and word consciousness.

However, Osa was disappointed and discouraged by the outcomes of the yearly standardized assessment used by her district, the Iowa Test of Basic Skills (ITBS). Her students' scores on the vocabulary subtest did not indicate any significant gains from their previous year's scores.

She knew that her students had increased knowledge about words, but she wanted quantitative evidence of that increased knowledge. If the standardized test scores did not demonstrate growth, was this instruction worth the time invested? What might be other evidence-based ways to document her students' growth?

"But the Words I Taught Weren't on the Test"

Vocabulary instruction plays an essential role during both literacy and disciplinary area instruction. Vocabulary knowledge is inextricably linked to reading comprehension and conceptual knowledge (Anderson & Freebody, 1985). The content disciplines are particularly rich areas for vocabulary development. Beck, McKeown, and Kucan (2002) referred to disciplinary vocabulary for which the concept is unknown as tier 3 words. Teaching tier 3 vocabulary requires situating the word within a system of ideas to be developed (Stahl & Nagy, 2006).

One of the challenges of teaching disciplinary vocabulary effectively is the paucity of available, classroom-friendly vocabulary assessments that can be used to inform instruction and to measure vocabulary growth, especially with the fastest growing sector of the school-age population—ELLs (National Clearinghouse for English Language Acquisition, 2007).

Often vocabulary is assessed at the end of a unit using a multiple-choice task, a fill-in-the-blank task or matching task. These modes of vocabulary assessment are shallow metrics of possible word knowledge. Further, more general measures such as the Peabody Picture Vocabulary Test (PPVT-III) or large-scale standardized tests that are used to compare students' vocabulary scores with a psychometrically derived norm are not helpful in informing instruction or sensitive to students' knowledge of lexical nuances.

What are some ways that we can gauge vocabulary development in the content areas? In this article, we articulate how the intricacies of word knowledge make assessment difficult, particularly with

The Reading Teacher, *63*(7), pp. 566–578
DOI:10.1598/RT.63.7.4
© 2010 International Reading Association
ISSN: 0034-0561 print / 1936-2714 online

disciplinary vocabulary. Next we address some considerations in improving teacher-made vocabulary tests and evaluating commercially produced tests.

We introduce a collection of techniques that teachers can adapt to provide evidence of vocabulary knowledge and vocabulary growth in the content areas that are appropriate for English-Only (EO) students and ELLs. We close with final thoughts for Osa and other teachers to encourage the development of contemporary content area vocabulary assessments that more precisely track students' vocabulary growth across the curriculum.

The Intricacies of Word Knowledge

Theoretical Underpinnings

The report of the National Reading Panel (NRP; National Institute of Child Health and Human Development [NICHD], 2000) and the implementation of No Child Left Behind resulted in an emphasis on the five pillars of reading: phonemic awareness, phonics, fluency, vocabulary, and comprehension. This emphasis has included a push to measure students' growth in each of these areas.

Commercially produced assessments of phonemic awareness, phonics, and fluency have proliferated. However, it is more challenging to find vocabulary and comprehension assessments that adhere to a conceptually rich construct that can serve as an instructional compass. This might be explained by Paris's (2005) interpretation of the five pillars within a developmental frame. Phonemic awareness, phonics, and fluency are considered *constrained* because they are fairly linear and students develop mastery levels (test ceilings) within a few years.

Alternatively, vocabulary and comprehension are multidimensional, incremental, context dependent, and develop across a lifetime. As a result, they simply do not lend themselves to simplistic, singular measures (NICHD, 2000; Paris, 2005). Our discussion addresses the unconstrained nature of vocabulary knowledge and describes some assessments that are suited to a complex theoretical construct.

What Does It Mean to Know a Word?

Knowing a word involves more than knowing a word's definition (Johnson & Pearson, 1984; Nagy & Scott, 2000). Word knowledge is multifaceted and can be characterized in various ways. Some facets of this complexity include (a) incrementality, (b) multidimensionality, and (c) receptive/productive duality.

Knowing a word is not an all-or-nothing phenomenon. Word learning happens incrementally; with each additional encounter with a word, depth of understanding accrues. Dale (1965) posited the existence of (at least) four incremental stages of word knowledge:

- *Stage 1*—Never having seen the term before
- *Stage 2*—Knowing there is such a word, but not knowing what it means
- *Stage 3*—Having context-bound and vague knowledge of the word's meaning
- *Stage 4*—Knowing the word well and remembering it

The final stage of Dale's conceptualization of word knowledge can be further broken down into additional stages, including the ability to name other words related to the word under study and knowing precise versus general word knowledge.

Instead of stages, Beck, McKeown, and Omanson, (1987) referred to a person's word knowledge as falling along a continuum. These include (a) no knowledge of the term, (b) general understanding, (c) narrow but context-bound understanding, such as knowing that *discriminate* means to pay special attention to subtle differences and exercise judgment about people but unable to recognize that the term could also be used to refer to singling out sounds in phonemic awareness activities, (d) having knowledge of a word but not being able to recall it readily enough to use it appropriately, and (e) decontextualized knowledge of a word's meaning, its relationship to other words, and extensions to metaphorical uses.

Bravo and Cervetti (2008) posited a similar continuum for content area vocabulary. These points on a continuum can range from having no control of a word (where students have never seen or heard the word) to passive control (where students can decode the term and provide a synonym or basic definition) and finally active control (where students can decode the word, provide a definition, situate it in connection to other words in the discipline, and use it in their oral and written communications).

For example, some students may have never heard the term *observe* while others may have a general

gist or passive control of the term and be able to mention its synonym *see*. Yet others may have active control and be able to recognize that to observe in science means to use any of the five senses to gather information and these students would be able to use the term correctly in both oral and written form. Such active control exemplifies the kind of contextual and relational understanding that characterizes conceptual understanding.

Word knowledge is a matter of degree and can grow over time. Incremental knowledge of a word occurs with multiple exposures in meaningful contexts.

> For each exposure, the child learns a little about the word, until the child develops a full and flexible knowledge about the word's meaning. This will include definitional aspects, such as the category to which it belongs and how it differs from other members of the category.... It will also contain information about the various context in which the word was found, and how the meaning differed in the different contexts. (Stahl & Stahl, 2004, p. 63)

Along the stages and continuum put forth by Beck et al. (1987), Bravo and Cervetti (2008), and Dale (1965), respectively, there are also qualitative dimensions of word knowledge. *Multidimensionality* aspects of word knowledge can include precise usage of the term, fluent access, and appreciation of metaphorical use of the term (Calfee & Drum, 1986).

Understanding that a term has more than one meaning and understanding those meanings is yet another dimension of word knowledge. Multiple meaning words abound in the English language. Johnson, Moe, and Baumann (1983) found that among the identified 9,000 critical vocabulary words for elementary-grade students, 70% were polysemous, or had more than one meaning.

Within content areas, polysemous words such as *property*, *operation*, and *current* often carry an everyday meaning and a more specialized meaning within the discipline. Understanding the shades of meanings of multimeaning words involves a certain depth of knowledge of that word.

Additional dimensions of word knowledge include lexical organization, which is the consideration of the relationship a word might have with other words (Johnson & Pearson, 1984; Nagy & Scott, 2000, Qian, 2002). Students' grasp of one word is linked to their knowledge of other words. In fact, learning the vocabulary of a discipline should be thought of as learning about the interconnectedness of ideas and concepts indexed by words. Cronbach (1942) encapsulated many of these dimensions, including the following:

- *Generalization*—The ability to define a word
- *Application*—Selecting an appropriate use of the word
- *Breadth*—Knowledge of multiple meanings of the word
- *Precision*—The ability to apply a term correctly to all situations
- *Availability*—The ability to use the word productively

Cronbach's (1942) final dimension leads us into the last facet of word knowledge, the receptive/productive duality. Receptive vocabulary refers to words students understand when they read or hear them. Productive vocabulary, on the other hand, refers to the words students can use correctly when talking or writing. Lexical competence for many develops from receptive to productive stages of vocabulary knowledge.

Vocabulary knowledge is multifaceted. Word knowledge is acquired incrementally. At each stage or point on a continuum of word knowledge, students might be familiar with the term, know words related to the term, or have flexibility with using it in both written and oral form. It is clear that to know a word is more than to know its definition. Teaching and testing definitions of words looks much different than contemporary approaches to instruction and assessment that consider incrementality, multidimensionality, and the students' level of use.

Vocabulary Assessment Considerations

Approaches to Vocabulary Assessment

Assessments may emphasize the measurement of vocabulary breadth or vocabulary depth. As defined by Anderson and Freebody (1981), *vocabulary breadth*

refers to the quantity of words for which students may have some level of knowledge. Multiple-choice tests at the end of units or standardized tests tend to measure breadth only. The breadth of the test itself may be extremely selective if it is testing only the knowledge of words from a particular story, a science unit, or some passive understanding of the word like a basic definition or synonym.

Furthermore, the breadth of the test is wider if testing students' knowledge of words learned across the year in all science units, for example, as might be found in a mandated state standardized test. However, even this is less comprehensive than a test like the PPVT-III or the ITBS, tests that choose a sample of words from a wide corpus. *Vocabulary depth* refers to how much students know about a word and the dimensions of word learning addressed previously.

Assessment Dimensions

As with any test, it is important to determine whether the vocabulary test's purpose is in alignment with each stakeholder's purpose. It is likely that this is the reason that Osa felt frustrated. The primary purpose of the ITBS is to look at group trends. Although it provides insights about students' receptive vocabulary compared with a group norm, it cannot be used to assess students' depth of knowledge about a specific disciplinary word corpus or to measure a students' ability to use vocabulary in productive ways.

In other words, current standardized measures are not suited to teachers' purpose of planning instruction or monitoring students' disciplinary vocabulary growth in both receptive and productive ways, or in a manner to capture the various multifaceted aspects of knowing a word (e.g., polysemy, interrelatedness, categorization; NICHD, 2000).

Read (2000) developed three continua for designing and evaluating vocabulary assessments. His work is based on an evaluation of vocabulary assessments for ELLs, but the three assessment dimensions are relevant to all vocabulary assessments. These assessment dimensions can be helpful to teachers in evaluating the purposes and usefulness of commercial assessments or in designing their own measures.

Discrete–Embedded. At the discrete end of the continuum, we have vocabulary treated as a separate subtest or isolated set of words distinct from each word's role within a larger construct of comprehension, composition, or conceptual application. Alternatively, a purely embedded measure would look at how students operationalize vocabulary in a holistic context and a vocabulary scale might be one measure of the larger construct.

For example, Blachowicz and Fisher's (2006) description of anecdotal record keeping is an example of an embedded measure. Throughout a content unit, a teacher keeps notes on vocabulary use by the students. Those notes are then transferred to a checklist that documents whether students applied the word in discussion, writing, or on a test. See Table 1 for a sample teacher checklist of geometry terms.

Even if words are presented in context, measures can be considered discrete measures if they are not using the vocabulary as part of a larger disciplinary knowledge construct. The 2009 National Assessment of Educational Progress (NAEP) framework assumes an embedded approach (National Assessment Governing Board [NAGB], 2009). Vocabulary items are interspersed among the comprehension items and viewed as part of the comprehension construct, but a vocabulary subtest score is also reported.

Selective–Comprehensive. The smaller the set of words from which the test sample is drawn, the more selective the test. If testing the vocabulary words from one story, assessment is at the selective end of the continuum. However, tests such as the ITBS select from a larger corpus of general vocabulary and are considered to be at the comprehensive end of this continuum.

In between and closer to the selective end would be a basal unit test or a disciplinary unit test. Further along the continuum toward comprehensive would be the vocabulary component of a state criterion referenced test in a single discipline.

Context-Independent–Context-Dependent. In its extreme form, context-independent tests simply present a word as an isolated element. However, this dimension has more to do with the need to engage with context to derive a meaning than simply how the word is presented. In multiple-choice measures that are context-dependent, all choices represent a possible definition of the word. Students need to identify the correct definition reflecting the word's use in a particular text passage.

Table 1
Teacher Checklist of Geometry Terms

Concept vocabulary	Javier	Darius	Ashley	Jade	Seann
Symmetry	D A T	D A T	D A T	D A T	D A T
Parallel	D W A T	D W A T	D W A T	A T	D W A T
Perpendicular	D A T	D A T	D A T	D A T	D A T
Congruent	D A T	D A T	D A T	D A T	D A T
Obtuse angle	D A T	A	D A T	D A T	T
Acute angle	D A T	A	D A T	D A T	T
Perimeter	D W A T	D W A T	D W A T	D W A T	D W A T
Area	D W A T	D W A T	D W A T	D W A T	D W A T
Inquiry vocabulary					
Figure	D W A T	D W A T	D W A T	D W A T	D W A T
Edge	D W A T	D W A T	D W A T	D W A T	D W A T
Translation	D A T	D A T	A	D A T	D A T

D = used in discussion; W = used in writing; A = applied; T = tested

Typically, embedded measures require the student to apply the word appropriately for the embedded context. Test designers for the 2009 NAEP were deliberate in selecting polysemous items and constructing distractors that reflect alternative meanings for each assessed word (NAGB, 2009).

Three Classroom Assessments

We intend that our selected assessments be used as a pretest and posttest providing a means of informing instruction as well as documenting vocabulary development during a relatively limited instructional time frame. There is empirical support for all three tasks (Bravo, Cervetti, Hiebert & Pearson, 2008; Stahl, 2008; Wesche & Paribakht, 1996). These studies applied the assessment to content area vocabulary, but each may be adapted to conceptual vocabulary within a literature theme. They are all appropriate for use with EO students and ELLs. Table 2 categorizes each assessment using Qian's (2002) Vocabulary Knowledge Dimensions and Read's (2000) Assessment Dimensions.

Vocabulary Knowledge Scale

The Vocabulary Knowledge Scale (VKS) is a self-report assessment that is consistent with Dale's (1965) incremental stages of word learning. Wesche and Paribakht (1996) applied the VKS with ELL students in a university course. They found that the instrument was useful in reflecting shifts on a self-report scale and sensitive enough to quantify incremental word knowledge gains.

The VKS is not designed to tap sophisticated knowledge or lexical nuances of a word in multiple contexts. It combines students' self-reported knowledge of a word in combination with a constructed response demonstrating knowledge of each target word. Students identify their level of knowledge about each teacher-selected word. The VKS format and scoring guide fall into the following five categories:

1. I don't remember having seen this word before. (1 point)
2. I have seen this word before, but I don't think I know what it means. (2 points)

Table 2
Design Features of Highlighted Assessments

Assessments	Qian's (2002) vocabulary knowledge dimensions	Read's (2000) assessment dimensions	Read's (2000) assessment dimensions	Read's (2000) assessment dimensions
Vocabulary Knowledge Scale	Depth	Discrete	Selective	Context-independent
Vocabulary Recognition Task	Size, depth, lexical organization	Discrete	Selective	Context-dependent
Vocabulary Assessment Magazine	Size, depth, productive knowledge	Embedded	Comprehensive	Context-dependent

3. I have seen this word before, and I think it means _____. (Synonym or translation; 3 points)
4. I know this word. It means _____. (Synonym or translation; 4 points)
5. I can use this word in a sentence: _____. (If you do this section, please also do category 4; 5 points).

Any incorrect response in category 3 yields a score of 2 points for the total item even if the student attempted category 4 and category 5 unsuccessfully. If the sentence in category 5 demonstrates the correct meaning but the word is not used appropriately in the sentence context, a score of 3 is given. A score of 4 is given if the wrong grammatical form of the target word is used in the correct context. A score of 5 reflects semantically and grammatically correct use of the *target* word. The VKS is administered as a pretest before the text or unit is taught and then after instruction to assess growth.

One important finding of Wesche and Paribakht's (1996) study of the VKS was the high correlation between the students' self-report of word knowledge and the actual score for demonstrated knowledge of the word. Correlations of perceived knowledge and attained scores for four content area themes were all above .95. This should help alleviate concerns about incorporating measures of self-reported vocabulary knowledge.

In addition, Wesche and Paribakht (1996) tested reliability for the VKS in their study of ELLs with wide-ranging levels of proficiency using a test–retest format. Although we cannot generalize to other vocabulary knowledge rating scales, Wesche and Paribakht obtained a high test–retest correlation above .8. Such a tool can potentially account for the confounding factors of many vocabulary measures, including literacy dependency and cultural bias.

It is possible to modify the VKS to assess the key vocabulary in content area units in elementary classrooms for even the youngest students. Blachowicz and Fisher (2006) applied the principles of the VKS in a table format—making it possible to assess a larger number of words. Kay (first author) used the Native American Home VKS (see Figure 1) as a pretest with her second-grade class. As a posttest, she used the VKS in conjunction with Figure 2, which required students to specify the tribe and resource materials used to build the home and to compose an illustration of the home.

Vocabulary Recognition Task

The Vocabulary Recognition Task (VRT) is a teacher-constructed yes–no task used to estimate vocabulary recognition in a content area (Stahl, 2008). Like the VKS, it combines self-report with demonstrated knowledge. Stahl applied the VRT with second graders reading at a mid–first-grade level. The purpose was to identify content-related words that the students could both read and associate with a unit of study.

In the study, each VRT consisted of a list of 25 words; 18 of the words were related to the content in

Figure 1
Native American Home VKS Pretest

Vocabulary words	I have never heard of this Native American dwelling.	I have heard of this kind of home, but I can't tell you much about it.	I can tell you what this home looks like and the materials used to make it.
Wigwam			
Apartment			
Longhouse			
Tipi/teepee			
Brush lodge			
Asi			

Figure 2
Native American Home VKS Posttest

Vocabulary words	Name of tribe who once lived in this home	Resources used to make this home	Draw a quick picture of this home
Wigwam			
Apartment			
Longhouse			
Tipi/teepee			
Brush lodge			
Asi			

each of four themed science units and 7 words were unrelated foils. See Figure 3 for a sample VRT comprised of vocabulary associated with the insect unit. Students circled the words that they were able to read and that were related to the topic. As a posttest, students completed the VRT and categorized the selected words under provided headings on a concept web (see Figure 4).

Anderson and Freebody's (1983) correction formula was applied to obtain a score that adjusts for possible guessing. A student scored a "hit" (H) when the word was circled correctly or a "false alarm" (FA) if an unrelated word was incorrectly circled. The proportion of words truly known, P(K), was determined with the following formula:

$$P(K) = P(H) - P(FA) / 1 - P(FA)$$

Webs received two scores, (1) total number of words correctly sorted by category and (2) percentage of words correctly selected on VRT that were correctly sorted by category.

The VRT requires teachers to select a bank of words that students are held accountable for in a content unit, thus measuring breadth of vocabulary knowledge on a topic. Using correlations with other vocabulary tests, Anderson and Freebody (1983) determined that the yes–no task is a reliable and valid measure of vocabulary assessment. They found that it provides a better measure of student knowledge than a multiple-choice task, particularly for younger students.

Teachers of novice readers know how important it is for them to be able to independently read words encountered in content units, something taken for

Figure 3
VRT

VRT - Insects
(Pretest and Posttest)

We *have been** learning about insects. Below you see a list of words. Put a circle around the words that you are able to read and are sure have something to do with insects. Do not guess, because wrong answers will lower your score.

thorax ✓	worm	ant ✓
ponds ✓	feelers ✓	bones
abdomen ✓	legs ✓	hatch ✓
ears	antenna ✓	cockroach ✓
mosquito	larva	spider FA
swamps ✓	snail	lungs
teeth	beetle ✓	wing ✓
colony	backyard ✓	molt
	bugs ✓	

Correctly Chosen Targets or Hits (H) 14

Incorrectly Chosen Foils or False Alarms (FA) 1

Known (K) word scores may be calculated using one of the following three options.

Classroom Score Calculation Option A (H – FA = K):
14 - 1 = 13

Classroom Score Calculation Option B (Percentage of Correct Choices = K):
14 + 6 = 20 20/25 = 80%

Proportion (P) Known Calculation Option C (P [H] – P [FA] / 1- P [FA] = P [K]):

$$\frac{14/18 - 1/7}{1 - 1/7} = \frac{.778 - .143}{1 - .143} = \frac{.635}{.857} = .741$$

* For pretests, change the wording to *will be*.

Figure 4
VRT Concept Web

VRT Web (Posttest Only)

Types of Insects	Places Insects Live
beetle bugs ant cockroach spider	ponds swamps backyard
Insects' Body Parts	**Other Insect Words**
thorax abdomen feelers legs antenna wing	hatch

Words Correct 14/18

Percentage of correctly categorized Hits 100%

granted with older students. This assessment is more adaptable to a larger corpus of target words than the VKS. The web that is included as part of the posttest provides a lens for depth of knowledge and lexical organization (Qian, 2002). Its simplicity also makes it a user-friendly format for ELLs.

Kay also used the VRT regularly in her second-grade classroom. Because the social studies and

science units were more in depth than the mini-units in the research project (Stahl, 2008), the classroom VRT typically contained a total of 33 words: 25 hits and 8 foils. When using the VRT in the classroom, a simple scoring system was used, H – FA or the percentage of correct responses. For classroom use, the web score was simply the total number of words placed correctly in each category.

Using the VRT as a pretest allows teachers to determine which words are known and unknown. As a result, less instructional time can be devoted to known words while providing more intense instruction to less familiar vocabulary.

In addition to learning about the students' vocabulary growth, the VRT posttest can assess our teaching. An interesting first-year consequence was discovering that there were weak pockets of instruction. For example, at the conclusion of the state-mandated unit on Australia the students did very well webbing animals and geographic regions of Australia. However, most students had less success webbing people and foods associated with Australia. This was a clear indication that the instruction and materials on these subtopics needed bolstering.

Vocabulary Assessment Magazine

The Vocabulary Assessment Magazine (VAM) was originally created to measure students' science knowledge, comprehension strategy use, and reading comprehension of science texts. Incidentally, as the analysis of the findings from this measure began, we noted that second- and third-grade students were using the science words in their responses to open-ended questions and were doing so with higher frequency at posttest than pretest (see Bravo et al., 2008).

Students were not prompted to use the science vocabulary in their responses during the time they completed the VAM. The words noted in students' responses included two types of words: (1) science inquiry, words that describe aspects of scientific investigations such as *observe*, *evidence*, *investigate*, and *predict*, like those found in Beck et al.'s (2002) tier 2 words and (2) science concept (e.g., *organism*, *erosion*, *shoreline*, and *adaptation*), words that Beck et al. would categorize as tier 3 words. Tier 3 words require conceptual development within a disciplinary construct.

There are two main parts to the VAM. The first includes brief reading passages with open-ended literacy questions pertaining to the passage. The open-ended questions associated with the passage prompted students to use comprehension strategies (e.g., making predictions, posing questions, making inferences, summarizing) and text feature (use of illustrations) knowledge.

The second part of the assessment is made up largely of science knowledge items. In Figure 5, students are asked to draw and label two different types of roots and write a sentence about their drawings. Drawing and labeling are literacy practices germane to the scientific enterprise, and the reason for their presence in the VAM is to measure students' science knowledge.

Another item, from a physical science unit, prompts students to "draw and describe the steps you would take to design a new kind of ice cream using flavorings, milk, and sugar as the main ingredients." These item types lend themselves to students' usage of both science inquiry and science concepts terminology, as they describe both a process and a larger scientific concept.

The analysis of the 703 VAMs completed by second- and third-grade students involved a frequency of word use. Statistically significant results were found for EO students and ELLs in the sample. On average, students were using 2.76 more science vocabulary at posttest than pretest. Gauging students' depth of word knowledge was possible through this alternative vocabulary assessment that involved students with authentic literacy practices used by the scientific community.

Although our research analysis addressed students' vocabulary use in the short-answer and open-ended questions in response to short, unfamiliar texts only, classroom teachers might consider additional practical applications of this format to assess vocabulary knowledge. Vocabulary frequency counts might be performed on students' responses to open-ended or essay questions on more traditional pre- and post-unit tests.

In addition, teachers can consider honoring students' approximations of terminology, perhaps assigning partial credit for imprecise use. The reason for assigning partial credit even for approximate uses is that if we consider Cronbach's (1942) receptive/productive duality using a term, albeit incorrectly, is more than having receptive knowledge of the term

Figure 5
Vocabulary Assessment Magazine Items

Life in the Forest

This magazine is called *Life in the Forest*.

What do you think the magazine will be about?

_____

_____

Draw and label two different types of roots.
Write a sentence under your drawings to describe the two types of roots.

1.

_____

_____

2.

_____

_____

and it is through use that we sharpen our understanding of vocabulary. Second, because one aspect of the multidimensionality of vocabulary knowledge is interrelatedness, it would be useful to note which additional vocabulary students used in concert.

Important considerations when implementing a format like the VAM to measure vocabulary knowledge include (a) insuring student access to any texts that students are asked to read and respond to, (b) documenting both inquiry and core conceptual vocabulary, (c) assuring that students have ample opportunities to use these terms in their responses, and (d) focusing on a core set of vocabulary words that can be taught extensively and to the point where students feel confident using them in oral and written form. A final consideration, although not a part of the design of the original VAM, is prompting students to use the vocabulary of the content.

Implementing a Content Vocabulary Assessment System

We recommend that grade-level teams of teachers work together to identify a list of targeted conceptual

vocabulary and inquiry-process words for each disciplinary unit. This list should include words that are essential for understanding the conceptual ideas and engaging in disciplinary activities within the unit. They are likely to be words that students will be held accountable for on assessments driven by the state standards. The words are pretested before the unit, posted on a content area word wall, deliberately taught, used (by both student and teacher) multiple times throughout the unit, and posttested at the conclusion of the unit (Stahl & Nagy, 2006).

In keeping with NRP recommendations (NICHD, 2000), teachers should use multiple measures to capture the multidimensionality of students' vocabulary knowledge. One possible system might be to use a general measure such as the VRT consistently for several units and to supplement it with more in-depth measures specific to disciplinary vocabulary (e.g., VAM, VKS, checklists of students' word use in oral or written form) that could be strategically developed over time in a phased approach.

Where Do We Go From Here?

We hope that we have provided useful suggestions for guiding Osa and other classroom teachers to document their students' content area vocabulary development. The measures that we provided as examples are to be viewed as a starting point for creating one's own assessment system, with attention to theoretical considerations. These types of measures are in keeping with the NRP's determination that current standardized measures lack sensitivity, provide only a baseline measure of global vocabulary knowledge, and that in practice, teacher-generated instruments are recommended (NICHD, 2000).

Although researchers are working to improve standardized measures, teachers can feel confident in taking an assertive stance in developing vocabulary assessments based on their own curriculum needs. In the words of the NRP, "the more closely assessment matches the instructional context, the more appropriate the conclusions about the instruction will be" (NICHD, 2000, 4.26).

References

Anderson, R.C., & Freebody, P. (1981). Vocabulary knowledge. In J.T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77–117). Newark, DE: International Reading Association.

Anderson, R.C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutson (Ed.), *Advances in reading/language research* (pp. 231–256). Greenwich, CT: JAI.

Anderson, R.C., & Freebody, P. (1985). Vocabulary knowledge. In H. Singer & R.B. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd ed., pp. 343–371). Newark, DE: International Reading Association.

Beck, I.L., McKeown, M.G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford.

Beck, I.L., McKeown, M.G., & Omanson, R.C. (1987). The effects and uses of diverse vocabulary instructional techniques. In M.G. McKeown & M.E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 147–163). Hillsdale, NJ: Erlbaum.

Blachowicz, C.L.Z., & Fisher, P.J. (2006). *Teaching vocabulary in all classrooms* (3rd ed.). Upper Saddle River, NJ: Pearson Education.

Bravo, M.A., & Cervetti, G.N. (2008). Teaching vocabulary through text and experience in content areas. In A.E. Farstrup & S.J. Samuels (Eds.), *What research has to say about vocabulary instruction* (pp. 130–149). Newark, DE: International Reading Association.

Bravo, M.A., Cervetti, G.N., Hiebert, E.H., & Pearson, P.D. (2008). *From passive to active control of science vocabulary* (56th yearbook of the National Reading Conference, pp. 122–135). Chicago: National Reading Conference.

Calfee, R.C., & Drum, P. (1986). Research on teaching reading. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 804–849). New York: Macmillan.

Cronbach, L.J. (1942). Measuring knowledge of precise word meaning. *The Journal of Educational Research, 36*(7), 528–534.

Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English, 42*(8), 895–901.

Johnson, D.D., Moe, A.J., & Baumann, J.F. (1983). *The Ginn word book for teachers: A basic lexicon*. Boston: Ginn.

Johnson, D.D., & Pearson, P.D. (1984). *Teaching reading vocabulary*. New York: Holt, Rinehart and Winston.

Nagy, W.E., & Scott, J.A. (2000). Vocabulary Processing. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–274). Mahwah, NJ: Erlbaum.

National Assessment Governing Board. (2009). *Reading framework for the 2009 National Assessment of Educational Progress*. Retrieved December 29, 2009, from www.nagb.org/publications/frameworks/reading09.pdf

National Clearinghouse for English Language Acquisition. (2007). National Clearinghouse for English Language Acquisition Report: NCELA frequently asked questions. Washington, DC: U.S. Department of Education. Retrieved October 2, 2007, from www.ncela.gwu.edu/faqs

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. (NIH Publication No. 00-4769). Washington, DC: National Institute of Child Health and Human Development.

Paris, S.G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184–202. doi:10.1598/RRQ.40.2.3

Qian, D.D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52*(3), 513–536. doi:10.1111/1467-9922.00193