9

Comprehending Through Composing: Reflections on Reading Assessment Strategies

Robert Calfee
Roxanne Greitz Miller
University of California, Riverside

The assignment to comment on the four chapters in this section, which present interestingly distinctive perspectives on the conference topic, is both engaging and challenging. The engagement springs from the depth of the ideas, the challenge from the task of molding the ideas into a coherent image. To address this task, we rely on three relatively standard lenses, and introduce a fourth that is less typical. Three constructs spring from the conference focus and the recent history of comprehension assessment: comprehension, (re)construction, and assessment. The fourth theme, composition, reflects our recent research, but, somewhat to our surprise, also emerged during the conference. Toward the end of the conference, for instance, Dick Anderson suggested that researchers might consider shifting attention from reading comprehension to literacy comprehension. Such a move is consonant with our thinking about the issues, and meshes with the increasingly important concept of academic language (Fillmore-Wong & Snow, 2000). Our chapter begins with brief reflections on the four lenses, continues with comments on the four chapters, and concludes by illustrating our recent efforts to engineer the reading-writing connection.

FOUR LENSES

Discussions of comprehension assessment arouse memories of "Rashomon" (Kurosawa, 1950), the Japanese movie classic demonstrating that any given situation can be perceived and construed in different ways. The articles presented at this conference are no exception, with varying construals of the two key terms. There was also frequent reference to a (re)constructivist stance toward these terms, and to the Rand Report (Snow, 2002). "Constructivist" denotes the active connection between text and reader, "re" to the reflective, metacognitive overlay. Composition is our addition to the mix, suggesting that one of the most trustworthy indicators of comprehension is the individual's capacity to compose a response to a text. Note that we did not say "write," which entails the mechanics of print. We must forego the temptation to compare the origins of comprehend and compose, although a story is to be found in the morphology.

COMPREHENSION

To comprehend a message is to understand or "get" it—right? Hence, the simple model of reading, which proposes that once the young child is taught to translate print into sound, then the existing natural language system kicks in and the child has learned to read. In "Understanding and Comprehending," Freedman and Calfee (1984) described the contrast between natural and formal language, the former referring to the variety of existing language registers that young children bring to school, and the latter to the academic register that serves as a standard for communication in educational settings, business, and government, and other middle-class exercises. From this perspective, the acquisition of literacy is the acquisition of a formal register with specific features, including an emphasis on explicitness, coherence, and attention to stylistic conventions. The argument, with wellsprings in discussions of the impact of literacy on thinking (Goody, 1977), was that literacy instruction for modern times should influence the way in which students think and communicate.

All kindergartners can understand a variety of linguistic messages, including those found in printed texts. They vary considerably in their experiences with words, sentences, and discourse patterns. Nevertheless, almost all children respond to language in a casual and nonstrategic manner. To be sure, some children have learned something about the "school game" (Heath, 1983), and know when and how to answer questions about the obvious: "How did the wolf feel when he fell down the chimney into the boiling water?"

ASSESSMENT

A decade ago, this construct would have required further explanation; only at the end of the conference, however, did Wixson suggest that this term required attention. For this chapter, three elements capture the contrast between testing and assessment (Calfee & Hiebert, 1988): purpose, method, and context.

Purpose is captured in several contrasts, most notably the difference between summative and formative evaluations of student performance, between growth and accomplishment. A related distinction is information that can be used to guide instruction versus indicators that predict later performance. External authorities increasingly strive to move summative operations into the classroom, suggesting, for instance, that externally mandated tests serve for "diagnosis."

Method encompasses variations, such as recognition versus production, multiple-choice versus short (or long) answers. "Testing," with the emphasis on cost-effectiveness, moves toward less expensive alternatives, generally with caveats about appropriate limitations on the results of such exercises. "Classroom assessment"—and the conjoined term has a meaning of its own—typically emphasizes the validity of the information for instructional decision making, and when integrated with instruction need not entail substantial increases in time. To be sure, this comment assumes a decision-making approach to instruction, which is not especially common. The most recent version of a cost-effective test is DIBELS (Dynamic Indicators of Basic Early Literacy; Kaminski & Good, 1998), a 1-min "production" task, where the job is to do as many simple things as possible within 1 min, including oral reading.

Context refers to the situation that surrounds the collection of evidence about comprehension. The testing mindset envisions the individual working in isolation, and distrusts any product that comes through social interaction— "Whose work is it?" The complement is direct instruction, wherein the teacher treats the class as a collection of individual learners. A different view places students within problem-solving groups, where comprehension is part of the process needed to obtain and use information required to complete a project. The teacher's assessment roles are complex in such settings, both for supporting group activities and for gathering and evaluating information about individual student growth and accomplishment.

(RE)CONSTRUCTION

The notion of comprehension as a constructive activity has been with us for a while. More recently, the idea has been grown in two interrelated ways. First,

from Vygotskian sources comes the notion of the social dimension of constructivist activities. Second, the "meta" label speaks to the importance of reflection. For these purposes, we propose three significant elements in reconstructive comprehension: (a) passive versus active, in which the reader can take in the words and sentences as they appear, compiling a collection of "propositions," or can approach the task more strategically, formulating hypotheses and instantiating schemata; (b) part versus whole, in which comprehension spans a continuum from the memorization of textual details toward deeper engagement in the macrostructural ideas that constitute a well-written passage; and (c), absorb versus transform, which on the surface is related to the active-passive contrast, but with different implications. A reader can rework the ideas in a passage, call on previous experience, integrate, summarize, and expand, all without attempting any major reshaping of the ideas. Transforming entails the use of comprehension outcomes to achieve results that transcend the original activity. The construct appears most obviously when the task is to combine two or more passages to create an entirely different product.

COMPOSITION

The fourth lens might seem a natural extension of the three previous ones, but as an alternative lens, it brings distinctive features to bear on the issues. Moreover, current practice suggests that more is involved than a "natural extension." For various reasons, reading and writing have become largely disconnected in present practice (Nelson & Calfee, 1998).

Like comprehension, composition encompasses a variety of meanings: spoken versus written, natural versus formal, formative versus summative. Kindergartners can neither "read" nor "write," but they can comprehend and compose in either natural or formal language registers. The third grader's journal contains a rich array of written material, typically casual accounts of personal experiences. District or state tests include a written composition, often to a decontextualized prompt providing students with limited guidance about purpose and audience.

To expand on this lens, consider the following scenarios springing from contemporary practice. First is the contrast from kindergarten "show and tell" activities to the research paper required from the late elementary grades onward. These two tasks differ in "medium," in the sense that kindergartners do not need to write, but also in the reliance on comprehension, in the sense that kindergartners can rely on personal experience, whereas a research paper typically builds on external texts.

Second is the way in which external texts serve as the basis for a composition. Reproduction is the classic requirement; the student answers questions about propositional specifics. Summarization is a slightly higher-level task; write an abbreviated version of a passage presenting major elements in relation to one another. A variation on this theme is the integrated summary, in which two or more passages are combined. A third level is the critical review, wherein the writer is asked to analyze and evaluate one or more texts. Quite different from any of the previous is the task that occurs when a passage serves as a basis for transformation, which requires comprehension, but more importantly, calls on the writer to use the passage for the creation of a new product (Bereiter & Scardamalia, 1987; Flower & Hayes, 1984). The nature of comprehension required for such tasks would seem to go beyond the notions underlying most conference discussions.

A third point centers around the role of integrated reading–writing activities within the classroom. As noted earlier, standard practice separates these domains. Connecting them poses various challenges: curriculum schedules, textbook materials, and the management of time during and across days and weeks. On the other hand, project-based learning, another label for this concept, offers the potential to enhance student motivation and support broad-based transfer of skills and knowledge.

THE FOUR PAPERS

This segment summarizes the four chapters in this section, reviewing each contribution through the four lenses. Other issues emerged during our review: The lenses entail implications for the value assigned to various educational outcomes. The developmental dimension pervades the discussion; what can (and should) children learn as they progress through the elementary grades? How do we deal with individual differences? Should conceptual and practical models emphasize the mean or the variance?

van den Broek

This chapter directly addresses the two conference themes: comprehension and assessment. The comprehension focus is on the early grades and stories, using a causal-network model as the foundation. Assessment is broadly construed as the capacity to "retell, apply, identify theme, critically appraise," with varying emphasis on these four elements. A developmental theme pervades the chapter; what are the varied ways in which young children process textual information (spoken or written, but primarily narrative) as they move from kindergarten into the

mid-elementary grades? In particular, in what ways do the causal networks that appear to underlie student responses change across these early years?

The research team created an assessment instrument around three principles: sensitivity to developmental changes, reliance on speech rather than print, and examination of profiles as well as single scores. Practically speaking, children viewed audio and video presentations of two popular narratives, "Blinky Bill" and "Rugrats," and were interviewed about their memories. The students generated more complete and complex responses across the years, but the changes were quantitative more than qualitative: "Preschool children engage in very much the same comprehension processes as do their older counterparts" (van den Broek, this volume). For instance, at all ages, children were more likely to recall the central story elements. In addition, relative standings remained much the same across the years; children who displayed higher levels of performance in kindergarten were still at the top of the heap at the end of second grade, even after factoring in differences in decoding skill. In this sense, the assessment exhibited predictive validity.

This chapter touches all four "lenses," including an assessment of students' capacity to reconstruct the essential elements of a passage. The particulars of "Blinky Bill" and "Rugrats" notwithstanding, the model offers interesting views about how the teachers of young children might delve into their charges' understandings of engaging passages. The presentation relies mainly on statistics to portray what also offers a rich qualitative image. The team seemed somewhat disappointed at the lack of more clear-cut developmental transitions; the Piagetian search for stages has a fascination that will not wane.

Neglected in this search is the potential of schooling to influence language and thought. That is, rather than concluding that development proceeds quantitatively more than qualitatively, another interpretation is that contemporary instruction may not typically produce fundamental changes in how children comprehend. More specifically, those children who enter kindergarten whose experiences align with the academic register are more likely to benefit from standards-based curriculum offerings, whereas those lacking these experiences may puzzle over what is going on. This hypothesis would account for the lack of developmental changes and the high levels of predictive validity, the basis for the "Matthew" effect.

What if teachers were to instruct youngsters in the secrets of the causal model, introducing such arcane terms as *character*, *plot*, and *theme*, and leading students to explore the role of motivation in the evolution of a narrative? What if students were provided a toolbox that they could use to "unbuild" a passage, or to build their own works? To be sure, such a strategy might undermine predictive validity by opening the way for all children to do rather remarkable things

regardless of their background. The conference focused on assessment, and so instruction remained in the background.

Paris and Colleagues

This contribution extends the previous one by inquiring into genuine and spurious correlates of comprehension. The challenge at the outset, of course, is to establish a standard—will the "real" comprehension indexes please stand up? The authors rely on the existing literature, which depends for the most part on "conventional outcome measures such as IRIs [informal reading inventories] and SATs [standardized achievement tests]" (Paris, this volume). Although the chapter offers a substantial variety of meaty findings and observations, it finesses the challenge of establishing a standard. The bottom line seems to be that the task of conceptualizing and operationalizing comprehension remains in such a primitive state that it is relatively easy to identify serious problems, even in the absence of a clear-cut standard.

The chapter critiques two spurious correlates: fluency and alphabet knowledge. Fluency refers mostly to the 1-min samples like those found in the DIBELS (Kaminski & Good, 1998) technique, mentioned throughout the conference. The researchers do not question the correlations; in a variety of settings, fluency correlates with comprehension. To be sure, the magnitude of the correlations depends on the particular comprehension measure and the developmental-achievement level of the students. Performance on surface-level measures (multiple-choice and cloze tests) are more likely to match with fluency indicators (interestingly, the "stimulus" does not seem to matter very much), and the relation is strongest in the earliest developmental stages, diminishing substantially by the midelementary grades.

What do the correlations mean, and how should they be used? The chapter reminds that correlation is not causality, and hence the researchers question the validity of claims that such instruments "assess" comprehension. Perhaps more significant are cautions about the instructional implications based on performance on such measures. The third grader who struggles to read a word list is probably going to have trouble comprehending a complex passage. Is the remedy to teach the student to read word lists more quickly? Perhaps not.

Alphabet knowledge offers another perspective on similar issues, with some additional fillips. For more than a half-century, research has shown a strong and persistent correlation between an entering kindergartner's knowledge of the ABCs, measured in various ways, and reading performance in later elementary grades, measured in a variety of ways. The chapter argues that the correlation is spurious for several reasons. For instance, although the correlation is strong

when ABC knowledge is measured on school entry, the relation quickly diminishes when ABC knowledge is measured later, for obvious reasons—by the end of kindergarten, virtually all children have been taught their ABCs.

"Spurious" has a negative connotation, from the Latin *spurius*; false or illegitimate. Is the kindergarten teacher mistaken in assessing alphabet knowledge and acting on it? Probably not, depending on the action and the interpretive basis for action. Preschool children acquire the ABCs for a variety of reasons, which, in combination, make alphabet knowledge a useful proxy for previous experience. Kindergartners are likely to learn their ABCs because they will be taught. However, the previous experiences continue to impact children's capacity to benefit from instructional activities, and in this sense, the initial assessment might provide useful information. The initial assessment can be predictive much like a blood pressure reading; a high reading calls for action. The pressure can be reduced in a variety of ways, but some are more effective than others in addressing the more fundamental problem. Likewise, teaching the ABCs is probably a good thing, but a "letter a week" is not necessarily the most effective way to introduce kindergartners to the full range of academic language competence.

Another interesting facet of ABC knowledge on kindergarten entry is statistical in character. The chapter notes that the distribution of ABC scores is seldom normal, in the sense of following the typical bell-shaped curve. In fact, at any given time a child either knows most or virtually none of his or her ABCs, producing a bimodal distribution, which actually enhances the potential of this simple indicator for decision making (Calfee, 1976). Spuriousness springs from misinterpretation and overuse.

Now to the lenses: this chapter addresses both comprehension and assessment, primarily focusing on methodological issues. At the end, the authors hint at some "genuine correlates"—language skills, receptive and productive vocabulary skills, and narrative reasoning, all pointing to the importance of construction and composition. They mention assessment procedures that might offer greater insight into underlying processes that would enhance both screening and diagnosis, which would seem to open the way for constructivist and compositional elements.

Stahl and Hiebert

"In the beginning was [and is] the word." Comprehension implies a passage, a collection of words, including the complex relations among these words—actually, among the constructs that they represent. What if the process stalls or fails at the word level? This question is of central concern in this chapter.

How does one think about the "word" as a starting point? At one level, translating a string of letters into a spoken response is important; "word recognition"

is one label for this construct. At a second level, connecting a string of letters to a semantic network is equally important; "word meaning" may reflect this assemblage. To be sure, these two interpretations of "word" carry quite different implications, cognitively and instructionally. This chapter starts with the "simple model," in which reading rests on decoding and oral language, with decoding essentially a word-level task. However, decoding without meaning is unlikely to promote comprehension, and so the authors explore three facets of "oral word reading": accuracy, rate, and prosody.

The authors' review of the literature in these areas is thorough and helpful, not because it resolves all of the issues, but by bringing attention to the relation between oral reading performance and comprehension. In brief, the conclusions are that (a) beyond third grade, accuracy does not seem sufficient to guarantee comprehension; (b) reading rate is correlated with comprehension on standardized measures, perhaps reflecting the impact of processing automaticity; and (c) prosody is difficult to pin down, is potentially important, and is not well researched.

Now to the lenses—at the outset, the authors' insistence that comprehension include a semantic component meshes with other contributions in this section. Their success in conceptualizing and operationalizing this component is less clear-cut, especially in the assessment arena. How might one design a vocabulary component to a comprehension exercise that illuminates specific and generic contributions at the "word" level? Oral reading in the early grades is common practice toward this end, and offers some insights into "word recognition." But which facets of this complex array of tasks best reveal the semantic and comprehension elements as separable entities, and in combination with one another?

This chapter does not directly address either constructive or compositional issues, but both offer openings to the role of word knowledge in comprehension assessment. We argue later for the critical importance of establishing the semantic basis for compositional activities. More to the point, it is probably unreasonable to ask anyone to "write" without an explicit textual base, either a specific passage or a well-defined set of experiences. In either instance, the resource will include words in one form or another. At a practical level, students are best positioned to compose when they have ready access to a collection of words, ideas, concepts, and relations. Assessments that do not provide this undergirding are likely to underestimate student competence.

Guthrie and Wigfield

The introduction of motivation into this section fills an often-overlooked gap. Paris suggested some time ago that achievement depended on both "skill" and

"will" (Paris & Oka, 1986), and we would add "thrill" to the list. The impact of the "age of accountability" appears in a parent's comment during a PTA meeting on reading programs: "I want my kid to learn to read, and I don't care if he wants to or not." Guthrie and Wigfield (this volume) focus on assessment more than comprehension, hypothesizing that low performance levels may reflect lack of effort more than competence. Their comprehension model includes several key facets, including background knowledge, text structure, causal networks, and integrative schema; the motivation model also covers several bases, including task mastery, intrinsic motivation, self efficacy, personal interest, and transactional beliefs. The review of relevant studies, including an investigation by the authors, shows that many facets in the two-part model correlate with one another. The most productive part of the chapter inquires into the influence on performance of situational characteristics, including both task conditions and inherent interest. Again, most studies are correlational, but the possibilities for experimental investigations appear obvious. A concluding research agenda suggests exploration of "relatively long, complex tasks ... [with] open-ended response formats that require extended writing" (Guthrie & Wigfield, p. 206, this volume) as situations that are more likely than brief decontextualized multiple-choice or short-answer tasks to engage and motivate students. In addition, the recommendation is routinely to inquire of students how interesting they found the task, and how hard they tried to do well.

Looking at the chapter through the four lenses speaks to the importance of motivational elements as essential components of effective comprehension assessment. The domain is largely ignored in policy and practice, and the chapter reveals the rather thin research base for making substantial claims. The (re)constructivist perspective offers conceptual (or at least metaphorical) support for considering motivation as part of the equation. Building is hard work, and will garner students' best efforts only when they are either pressured (a frequent strategy) or intrinsically engaged (relatively rare, especially for students with reading problems). Long, complex tasks requiring extended writing would seem to pose unwelcome challenges to many students, further hindering genuine engagement. In the next section, we describe a scenario that addresses some of these issues.

A motivational ingredient not mentioned by Guthrie and Wigfield (this volume) is the social context for the comprehension-composition task. As noted earlier, assessment often carries the connotation of individual efforts, and group tasks are suspect. To be sure, most "outside and beyond school" tasks involve cooperative activities, and techniques are available for sorting out individual and collective contributions. We do not attempt to review the literature on the mo-

tivational concomitants of social versus group activities, but the positive consequences of working together would seem to warrant further consideration.

A Conceptual and Practical Example

Constructing a conceptual framework that joins comprehension, composition, and assessment, and then translates this concept into successful practice, has been a major focus of our work over recent years (Calfee & Miller, 2003; Miller & Calfee, 2004a; Miller & Calfee, 2004b). The basis for the framework rests on the Vygotskian theories already mentioned, and relies extensively on the constructs of "schema theory" (Anderson, Spiro, & Anderson, 1978). The schema construct provides a unifying framework for linking comprehension, composition, and assessment; understanding a text requires connecting with an existing memory framework that contains "slots" for incoming information, and that establishes prospective relations among these elements. A similar conceptualization applies to composition; the author chooses a framework to guide the assembly of known and new elements during composition.

Schema theory applies with particular force to the design of constructivist assessment tasks. The challenge here is to tap into both processes and products as students reconstruct and "transconstruct" textual materials. The challenge is to delve into students' thinking to evaluate and—more importantly—to shape their capacities to work with ideas and communicate with others. Elsewhere we have presented the "Read–Write Cycle" as a conceptual bridge for connecting schema constructs with the practicalities of the "research paper" assignment, the classroom commonplace where comprehension, composition, and assessment are most frequently juxtaposed (see Calfee, 1998; Miller & Calfee, 2004a). We conclude this review with a practical analogue to the Read–Write Cycle, which we offer partly as a concrete example of the potential for valid reading comprehension assessment through appropriately designed composition activities, and partly to suggest the potential for enhancing composition assessment through appropriately designed comprehension activities. Moreover, this combination provides a model for effective integrated literacy activities across a broad range of content areas, and across the span of formative and summative assessments.

The example, which might seem mischievous on our part, builds on CLAS, the California Learning Assessment System, which for 2 years served as the primary vehicle for California's statewide evaluation of literacy achievement. Following a rough start because of implementation shortcomings, CLAS almost immediately captured the attention and commitment of classroom teachers across the state as the type of assessment that warranted their invest-

ment; it was a test worth teaching to. The demise of CLAS is a story in its own right, reflecting a variety of concerns including (a) the feasibility of performance-based assessment, (b) technical concerns about reliability and SEM (standard error of measurement), and (c) choices of reading materials and performance tasks (Underwood, 1999). CLAS-Plus builds on the constructivist framework that undergirded CLAS, with modifications reflecting the opportunities available in a nonstandardized environment. As shown in Fig. 9.1, students (a) read and respond to a passage, (b) meet in small groups to discuss their responses in preparation for (c) a writing composition based on the text. The augmentations in CLAS-Plus include (a) introductory scaffolding of the topic, facilitated by the teacher through "webbing" activities; (b) posting of the products of student discussions throughout the classroom (practically speaking, lots of words presented on public display); (c) provision of graphic organizers for both comprehension and composition; (d) explicit discussion of performance criteria or rubrics; and (e) inclusion of a social component of the activity, including project presentations at the end of the exercise. Teachers praised CLAS because it appeared to be valid, the tasks were interesting and engaging, and it offered a workable classroom model. The writing assignments (compositions) attended to audience, purpose, and voice. The cooperative phases, wherein students shared their knowledge and views on the topic, enhanced students' reflective and critical stance during their writing and enhanced motivation.
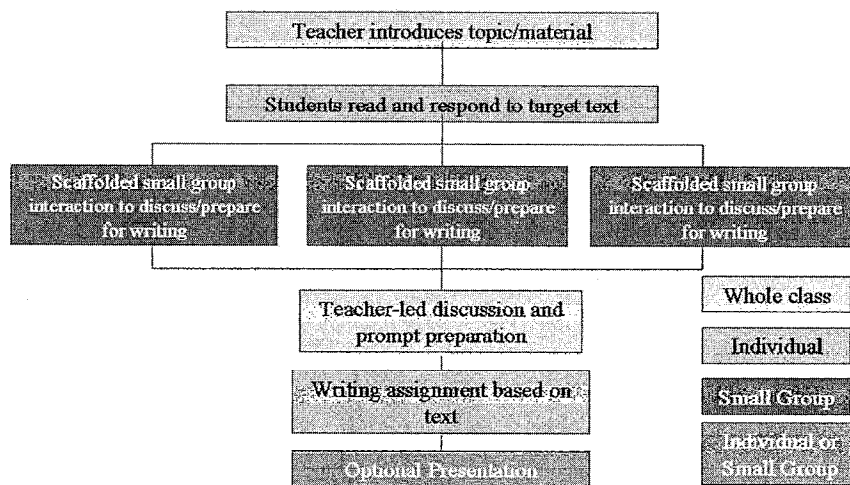

FIG. 9.1.   California Learning Assessment System-Plus Sequence of Activities (Calfee & Wilson, 2004). Reprinted with permission.

The "Reading and Writing About Science Project" (RWS) has employed the CLAS-Plus design as both an instructional and assessment model within the Read–Write Cycle framework. The aim in RWS was to evaluate the impact of content-area-embedded literacy activities to enhance mid-elementary grade students' comprehension of expository passages of the sort found in science, and to improve their competence in composing research reports. The project has been based in California, where the accountability emphasis in recent years has focused on reading and math—finding a niche for science in "low-performing" schools has proven a challenge.

During RWS, students were exposed to curriculum "blocks" based on a single science theme (e.g., the rock cycle or plate tectonics). Each block introduced three different reading samples using the CLAS-Plus format, following the sequence of steps illustrated in Fig. 9.1. The subsequent example from the rock cycle block illustrates how schema theory was interwoven through comprehension, composition, and assessment.

During an introductory lesson on the rock cycle, the teacher first identifies for students what will be studied (in this case, different kinds of rocks and how they are formed). Teachers activate students' prior topic knowledge (Alexander, Schallert, & Hare, 1991) and preexisting schema by having them actively reflect, share with others, and use prewriting and other reflective techniques as brainstorming methods (see Miller & Calfee, 2004a, 2004b). Students write down and share their knowledge and experience in whole class and small groups about different kinds of rocks and their origins, and make predictions about the content of the upcoming reading sample.

Students then (a) read a reading sample on the stages of the rock cycle (igneous, sedimentary, metamorphic), use "think-aloud" strategies while reading individually, and conduct analysis of text structure, purpose, and audience; (b) organize prereading and postreading concepts using graphical structures; and (c) use contextual clues in the text to translate new and unfamiliar vocabulary. Graphic organizers are not given to the students; instead, students, with teacher guidance, actively construct an organizer appropriate to the context, justifying their organization of the content matter into particular graphic structures. Defense of the organizer undergirds students' metacognitive and reasoning ability and engages them in creating the structure that works best for them (Chambliss & Calfee, 1998).

The think-aloud procedure (Davey, 1983), voicing and writing down thoughts as the text is read, either as teacher modeling or student self-monitoring, appears effective in raising students' reading comprehension. RWS teachers are encouraged to model think-aloud procedures with students prior to reading. As they read, students are instructed to write both their observations and questions onto the reading sample copies, and to monitor their own com-

prehension. Written comments from think-aloud exercises also serve as a bridge to the composition phase.

Vocabulary development through context clues is also incorporated in the read-and-respond portion of instruction. We agree with arguments that comprehension depends on word-level processing. Acquisition of context strategies for vocabulary development provides students a transferable method that applies to all subject areas (again, creating a "slot" for students to rely on across other multiple subject matters). In RWS, teachers developed vocabulary exercises from the assigned readings so that students derived word meanings from the text itself rather than simply looking up words in the dictionary. For example, *metamorphic* was a key term in the rock cycle unit (referring to both a rock type and a stage in the rock cycle). Many students had heard of *metamorphosis*, but only considered this term in relation to living things like caterpillars and butterflies. The application to describe changes in rocks was not obvious to them, and had to be explored in the full context of the target texts to reveal the meaning, and to construct new schemata.

After reading the text sample, students examine the structure and content of their graphic organizer, facilitated by the teacher. Students may discard, reorder, or restructure their ideas, which may be incorrect, inaccurate, or simply irrelevant. The costs of changes at this stage are relatively modest—nothing has been "written." Students share their reflections on the reading in small groups and with the teacher, again serving to further externalize and shape students' reflections on the content knowledge transmitted through the reading.

The teacher introduces the writing prompt and students proceed to reflect on the task. Writing prompts used for assessment in the Read–Write Cycle follow specific guidelines developed by Miller and Calfee (2004), which also teaches students to "dissect" the prompt into its constituent elements, to locate ideas from the reading, and to translate the information into a writing plan. The frequent use of defined "prompt elements" creates a schema for students to use when faced with an assessment task, extending schema theory to encompass the full assessment spectrum. Students understand the purpose of their writing, the intended audience for the writing, the form that the writing is to take, and the type of supporting details to use in their writing. When faced with subsequent assessments, they have access to a packet of methods for prompt and passage deconstruction and composition construction.

The final task is writing the individual compositions. The writing task provides an opportunity for students to synthesize, transform, and apply knowledge. This extension is performed individually, with no assistance from peers or the teacher. After composing, students share their compositions with peers in small groups or whole class interactions. Opportunities for students to liter-

ally "compare notes" expose them to different interpretations and points of view, and to varying levels of writing expertise. The "Writing to Models" approach is subtly at work here; good examples of student writing delivered to students by students provide a standard for future compositions. The final drafts are scored using multiple rubrics (Miller & Calfee, 2004a), reflecting both standard writing gauges (e.g., grammar, mechanics, vocabulary) and transmission of content knowledge. The specific attention to content knowledge makes this assessment strategy a more comprehensive representation of student comprehension.

FINAL THOUGHTS

Comprehension may arguably be viewed as a definitive cognitive achievement in its own right. Unfortunately, this accomplishment can be fully appreciated only when made public in some fashion. To be sure, the individual may experience great internal delight after struggling with a message and finally "getting it." But how can external observers (teachers and researchers) tap into this experience, assuming a good reason for such an attempt? The most direct and comprehensive approach is to ask the individual to present the results of the activity, by retelling, summarizing, applying, critiquing, extending, transforming, and so on—in brief, by composing some sort of response. A range of indirect tactics is also available: multiple-choice and short-answer queries, surveys of "feeling of knowing," and other less direct indicators.

In reviewing the chapters in this section, we have emphasized a direct approach in which comprehension is connected with composition, the latter a definitive cognitive achievement in its own right. We offer three arguments in support of this proposal. First, it builds on a defensible pedagogical model for promoting the growth of formal language and literacy. The model has a long history stretching back to the Greek rhetoricians; in today's world, it is increasingly important that these "secrets" to effective communication become available to all of our citizens.

Second, the read–write model embodied in the Read–Write Cycle and CLAS-Plus turns out to be practically workable in classroom settings, both for assessment and instruction within literacy programs, but also readily extendable to other content domains. Moreover, the model provides a foundation for teachers' professional development in literacy as the basis for integrated projects that support students' thinking and communication skills in the elementary grades. Rather amazingly, given various snafus in California's initial implementation of CLAS, teachers' memories of the program remain generally positive, rather uniformly evoking the response that CLAS tested what teachers ought to be teaching (Underwood, 1999).

The third point addresses the question of the validity of assessing comprehension through composition. The concern is that the composing task presents barriers to adequate assessment of student understanding; students may "know" much more than they can express in writing. Our response to such concerns is twofold. We would suggest that first, students may also "know" much more (or less) than is likely to be revealed by other means of assessment. This suggestion has instructional implications as well. The most effective tactics for enhancing performance on multiple-choice comprehension tests, for instance, have more to do with test taking than with passage comprehension. In the age of accountability, teachers are well advised to consider "what works" for the privileged indicators, which means teaching directly to the test. Our second claim, based on achievement patterns emerging from the RWS, is that effectively scaffolded reading–writing experiences enhance both comprehension and composition. Students taught that reading is reading and writing is writing are unlikely to be able to demonstrate either comprehension or composition skills in a CLAS-like situation; instructional experiences that integrate the two domains are essential to our argument.

Several counterarguments to the proposal also warrant consideration. The first is the pressure springing from the current accountability systems, which emphasize reading and efficient correlates of reading comprehension to the neglect of writing. Related to this point is the inertia embodied in instructional materials, which determine both curriculum and pedagogy, and which today privilege reading. The reading–writing model requires support that is generally not available in today's materials.

Second is the implicit assumption in our argument that the classroom teacher is capable of the professional judgments required to manage complex projects, which call for ongoing adaptations in the original instructional plan in response to emerging needs and opportunities. For the proposed model to work, control would spring more from classrooms than statehouses.

A third problem arises from "grain size"—over the past half-century, textbook publishers and test makers have created templates that emphasize the reading "lesson," a set of activities lasting for an hour or so, during which a series of objectives are covered by the teacher, in accord with the checklist format of the typical standards-based scope-and-sequence chart. Objectives are introduced, reviewed, and tested across a series of lessons. The integrity across lessons rests on a passage that students will encounter for a week or so, depending on holidays and other less predictable events. In the read–write model, integrity builds on a series of interrelated activities all aimed toward completion of a "construction" of substantial dimensions, a grain size measured more in weeks than in minutes. An aside—the challenges for instructional design arise not

from the concept of educational standards, but from the grain size used to create and implement the standards. In general, most national and state standards begin with laudable outcomes at the highest level of the design, but then are overwhelmed by the steady accumulation of lower-level objectives that serve as the operational basis for textbook and test materials.

Finally, the No Child Left Behind act emphasis on the "children left behind" undermines any suggestion that demanding tasks should be the actual basis for assessing the achievement of students from at-risk backgrounds. High standards are operationalized as (arbitrarily) high scores on low-level tests. Relying on complex projects that call for both comprehension and composition as a way of judging the achievements of students from at-risk backgrounds is an idea that evokes disbelief and derision in many quarters. And yet, that is our proposal.

How do we achieve the proposed outcomes, especially for those students most in need? Instruction would seem to be the answer. The conference and the chapters in this section centered on assessment, for understandable reasons. However, the integration of instruction and assessment makes sense both conceptually and practically. A quarter-century ago, Durkin (1978) suggested that something was amiss with comprehension instruction in the elementary grades. We are not aware of any recent reports demonstrating any substantial change in this state of affairs. We also have not encountered any parallel investigations of the situation for composition instruction. Perhaps worth consideration as a research agenda ...

ACKNOWLEDGMENTS

REFERENCES

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61,* 315–343.

Anderson, R. C., Spiro, R. J., & Anderson, M. C. (1978). Schemata as scaffolding for the representation of information in connected discourse. *American Educational Research Journal, 15,* 433–440

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Calfee, R. C. (1976). Sources of dependency in cognitive processes. In D. Klahr (Ed.), *Cognition and instruction: 10th Annual Carnegie-Mellon Symposium on Cognition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Calfee, R. C. (1998). Leading middle-grade students from reading to writing: Conceptual and practical aspects. In R. C. Calfee & N. N. Nelson (Eds.), *The reading–writing connection: The ninety-seventh yearbook of the National Society for the Study of Education* (pp. 203–228). Chicago: University of Chicago Press.

Calfee, R. C., & Hiebert, E. (1988). The teacher's role in using assessment to improve learning. In C. V. Bunderson (Ed.), *Assessment in the service of learning* (pp. 45–61). Princeton, NJ: Educational Testing Service.

Calfee, R. C., & Miller, R. G. (2003, April). *Embedding reading and writing instruction in the content area.* Paper presented at AERA National Convention, Chicago.

Calfee, R. C., & Wilson, K. M. (2004). A classroom-based writing assessment framework. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (p. 590). New York: Guilford.

Chambliss, M. J., & Calfee, R. C. (1998). *Textbooks for learning: Nurturing children's minds.* Malden, MA: Blackwell.

Davey, B. (1983). Think-aloud modeling cognitive process of reading comprehension. *Journal of Reading, 27,* 44–47.

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14,* 481–533.

Fillmore-Wong, L., & Snow, C. E. (2000). *What teachers need to know about language.* Washington, DC: Center for Applied Linguistics.

Flower, L. S., & Hayes, J. R. (1984). Images, plans, and prose: The representation of meaning in writing. *Written Communication, 1,* 120–160.

Freedman, S. W., & Calfee, R. C. (1984). Understanding and comprehending. *Written Communication, 1,* 459–490.

Goody, J. (1977). *The domestication of the savage mind.* Cambridge, England: Cambridge University Press.

Heath, S. B. (1983). *Ways with words.* Cambridge, England: Cambridge University Press.

Kurosawa, A. (Director). (1950). *Rashomon* [Motion picture]. Japan: RKO Pictures.

Kaminski, R. A., & Good, R. H., III. (1998). Assessing early literacy skills in a problem-solving model: Dynamic indicators of basic early literacy skills. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 113–142). New York: Guilford.

Miller, R. G., & Calfee, R. C. (2004a). Building a better reading/writing assessment: Bridging cognitive theory, instruction, and assessment. *English Leadership Quarterly, 26,* 6–13.

Miller, R. G., & Calfee, R. C. (2004b). *Breaking ground: Creating opportunities for authentic literacy instruction in the upper elementary content areas.* Manuscript submitted for publication.

Miller, R. G., & Calfee, R. C. (2004c). Making thinking visible: A method to encourage science writing in upper elementary grades. *Science and Childen, VV(N),* 20–25.

Nelson, N. N., & Calfee, R. C. (Eds.). (1998). *The reading–writing connection: The yearbook of the National Society for the Study of Education.* Chicago: University of Chicago Press.

Paris, S. G., & Oka, E. R. (1986). Self-regulated learning among exceptional children. *Exceptional Children Special Issue: Competence and Instruction, 53,* 103–108.

Snow, C. (2002). *Reading for understanding: Toward an R&D program for reading compre-hension*. Washington: RAND.
Underwood, T. (1999). *The portfolio project: A study of assessment, instruction, and middle school reform*. Urbana, IL: National Council of Teachers of English.