Language assessment and writing disorders

Sarah Warshauer Freedman, PhD
Assistant Professor
School of Education
University of California, Berkeley
Berkeley, California

ALTHOUGH THE 3Rs include reading, 'riting, and 'rithmetic, writing seems to have been the neglected basic skill. One of the foremost researchers on children's writing, Donald Graves (1978), has documented the lack of attention to the teaching of writing in the American elementary school. He attributes this lack of attention to several sources. For one, he points out that paradoxically, "the so-called return to basics vaults over writing to the skills of penmanship, vocabulary, spelling, and usage that are thought necessary to precede composition. So much time is devoted to blocking and tackling drills that there is often no time to play the real game, writing" (p. 10).

A second phenomenon that Graves notes is the inappropriate dominance of reading instruction: "Our anxiety about reading is a national neurosis. . . . Concern about reading is today such a political, economic, and social force in American education that an imbalance in forms of

0271-8294/82/0024-0034$2.00

communication is guaranteed from the start of a child's schooling" (p. 10).

Graves argues for a more balanced view of reading and writing in contemporary education, a view that he thinks can lead educators to capitalize on the contributions that learning one language skill can contribute to the learning of another. Gordon Wells (1981), a British researcher studying childrens' language development, has observed similar inattention to writing in England. Like Graves, Wells notes that "many discussions of literacy, particularly of the early stages, give most of their attention to reading" (p. 273). Wells goes on to say that

> it is in the process of composition . . . that one ultimately becomes most fully aware of the power—and limitations of the written language. . . . In my view . . . the common practice of waiting until children are already quite fluent readers before encouraging them to compose written text is to miss one of the most effective means of introducing them to the possibilities that this new medium of communication makes available.
>
> The reason usually given for this separation of receptive and productive activities is the difficulty children have in mastering the skills of handwriting. But this need not be a deterrent. . . . One possibility is for children to "write aloud" their stories, notices, observations and so on. (pp. 273–274)

Similarly, the literature on language-learning disabilities reflects this unbalanced attention to the mechanical aspects of writing and bias toward reading over writing. Moran (1981), in a recent study that compares the writing of learning-disabled and low-achieving secondary students, justifies her focus on measuring mostly mechanics by referring to the work of Howerton, Jacobson, and Selden (1977) who stress the following: "In evaluation of written expression, formal features have been emphasized over elements such as ideation, organization of ideas or support for arguments, which are acknowledged to be more important but also more difficult to assess objectively" (Moran, 1981, p. 271).

Moran continues her justification by claiming that teachers respond more to formal features than to the higher levels of the discourse. However, she lacks evidence for this claim. Recent work (Freedman, 1979) has shown that teachers respond first to the development and the organization of the ideas and that only after the student has demonstrated control over these higher levels of the discourse do teachers respond to the sentence structure and mechanics. Furthermore, most studies that contain an attempt to correlate countable features of students' writing with teachers' evaluations show that the best predictor of the evaluation is the length of the passage. Length is more likely an index of idea development than mechanical issues. Mechanical indices are usually counted in such studies (e.g., number of spelling errors, sentence or T-unit length), and such indices have not been found to contribute as much to the score as does length (e.g., Grobe, 1981; Nold & Freedman, 1977; Slotnick & Knapp, 1971). Both in teaching and assessment it seems critical that techniques be developed for examining all features of the written piece.

The bias toward reading over writing in the literature on learning disorders can be found even in Myklebust's (1965) treatise

on testing writing in which he stresses that reading acquisition must precede the acquisition of any kind of writing. He does not take into account the fact that when most children enter school, they already possess a great deal of skill in producing oral language and have spent many hours listening to written stories, observing others write, and perhaps engaging in writing activities as well. Like both Graves and Wells, Dyson (1980) suggests that children's knowledge about written language supports their early writing, and Loban (1976) and Dyson (1981) further demonstrate the influence of oral language on writing.

One might think that once children learn to read and once they have mastered the skills of handwriting the curriculum would begin to place more emphasis on writing. However, Applebee (1981) documents the fact that the amount of writing being practiced in the secondary school is limited. Although attempts are being made to infuse writing into the curriculum at all levels and across disciplines, most notably by the in-service teacher training network established by the University of California's Bay Area Writing Project and its many national and international affiliates, changes are not immediate, and the number of teachers reached by the Writing Projects, although impressive, is not comprehensive.

As educational leaders begin to argue for teaching writing and reading simultaneously and for teaching more writing at all grade levels and in all disciplines, the issues surrounding the assessment of writing and the identification of writing disorders come to the forefront of the educator's list of priorities. Once a subject becomes integral to the curriculum, it becomes important to identify children who have difficulty and provide them with special help. Lerner (1981) notes that writing is the most prevalent communication difficulty for children labeled as *learning disabled*. Poplin, Gray, Larsen, Banikowski, and Mehring (1980) demonstrate that writing problems become more severe as children get older.

If writing instruction and assessment are to be improved, tests for identifying and diagnosing the needs of children with writing problems must encompass several issues surrounding the assessment of student writing. Two existing instruments for testing those suspected of having writing disorders (Myklebust, 1965; Poteet, 1980) will be mentioned, and then some of the important issues facing writing assessors who hope to examine writing in depth will be discussed.

The best known test for assessing the writing of children with language-learning difficulties is Myklebust's (1965) Picture Story Language Test. The test, which uses a picture prompt to elicit a written narrative, is scored for total words, total sentences, words per sentence, syntax, and level of abstractness. Myklebust's picture prompts are somewhat outdated, and his scoring instructions do not go much beyond the superficial aspects of writing. Recently, Poteet (1980) has argued for informal assessments of writing and has offered a checklist of written expression skills. His checklist is comprehensive in the mechanical areas of penmanship, spelling, and grammar; however, it is sketchier on the level of ideation. In this last category, he includes syntax and word counts along with directions for measures

of comprehensibility, style, and tone. Poteet reminds the reader that the checklist is only as good as the teacher who administers it and that it is informal and cannot be substituted for normative data. Both Poteet and Myklebust provide helpful information about how to assess a given writing sample.

Interestingly, on the surface, it seems that it would be easy to detect writing problems. When children write, they leave a tangible product that can be assessed directly. The product either does or does not contain evidence of writing problems. But difficulties plague all of us concerned with writing assessment. Important issues that must be considered in developing criteria for comparing the writing of different students are (a) the consistency of the results the writer will produce on one occasion versus another, (b) the consistency of the judgments of different evaluators about the same piece of writing (setting standards and levels of expectation) and the choice of scoring methods, (c) the expense compounded with the comparability problems for direct measures of writing versus the validity problems of indirect measures, (d) the determination of the reasons for a writing problem once such a problem has been identified, and (e) the design of an effective plan for helping the student solve the problem.

COMPARABILITY OF RESULTS

Environment

All of us who write know that sometimes we produce better works than other times. With writing, unless we are performing a perfunctory, formulaic task, we often have difficulty writing on command. This happens because when writing we must feel that we have something to say, unlike when reading when we only must receive what someone else says to us. Just as some people may not wish to talk when we have nothing to say, they resist writing when we have nothing to say. Students can look disabled as speakers and as writers when they do not feel any need or urge to communicate. Older, schoolwise students learn to consider writing test essays a perfunctory task, and so may perform well without feeling a need to communicate. Younger or less savvy students may perform poorly on a test-writing task when they feel no need to communicate because they cannot make the task perfunctory. These same writers may perform capably in writing situations in which they are more motivated.

Just as the special environment of the writing test may inhibit the writing of some students, so too may the familiar classroom inhibit rather than facilitate writing (others have shown how the classroom can inhibit spoken language, especially of minority students; see, e.g., Cazden, John, & Hymes, 1972; Wells, 1981). Even if a tester evaluates the student's performance in the classroom and in a testing situation, these two observational

Even if a tester evaluates the student's performance in the classroom and in a testing situation, these two observational settings may converge to depict a disabled writer who may not be disabled.

settings may converge to depict a disabled writer who may not be disabled. This same writer in another classroom, in another test situation, or in some informal environment might not appear disabled.

Whatever the testing or writing context, it is important to provide a supportive testing environment that includes time for encouraging the child to engage in a full writing process, from idea generation through revision. Also important, as Myklebust (1965) noted, is the fact that writing is a skill that is learned in school. He warns that "the diagnostician must constantly be aware of the importance of the teaching methods to which a child has been exposed and must attempt to distinguish between those showing poor success in written language because of inadequate teaching and those who are deficient for other reasons" (p. 65).

Process

Classroom, testing, and teaching environments aside, variables during the composing process itself can conspire against a writer. I recently watched one fifth grader change from Jekyll to Hyde at different stages during the process of completing a single writing task. She began the task the day before it was due and seemed to panic because she had so little time to complete it. Instead of working productively, she spent much of her time saying, "I can't," or "I need help" (meaning "Will you do it for me?"), fretting, crying, and complaining alternately. I tried to help by stimulating her idea-generation process, but she rejected this help. By the end of this session, she had produced little, and what she had produced was inappropriate for her audience, difficult to understand, marred by mechanical errors, and sloppy in appearance.

The next day she tried to revise her draft. At this point, she decided that in spite of the short amount of time she had, she would have to begin again. With her new approach to the topic, she was able, on this new draft, to write an excellent paper. The changes in her writing took place at all levels of the discourse. Her spelling and punctuation were more accurate, her handwriting was more legible, and her ideas were better developed with sufficient appropriate detail to communicate to her reader. During this second drafting process, the writer did not ask me for help other than to sit with her and act as an instant dictionary for difficult-to-spell words. This support allowed her to produce her prose more rapidly and to write more fluently; she did not have to interrupt her flow of words.

The piece produced on the first day indicated a writer who should have been sent to a learning specialist for help with handwriting, spelling, punctuation, and idea development. The piece produced on the second day showed a competent fifth-grade writer, not in need of special assistance. On the first day, the writer had not decided how to approach the writing task, and she allowed the close deadline to unnerve her; on the second day she decided on an approach and used the even closer deadline to her advantage. This anecdote shows the potentially deceptive nature of the written product in the assessment of writing skills. It also may show how a single writer can vary greatly in level of performance from one attempt to the next.

Topic

Another anecdote illustrates how another variable within the writing situation, the topic, can significantly influence students' writing. I recently heard a college writing teacher criticize a topic she had assigned to her freshmen because it elicited good writing, even from her least able students. Commonly, unbeknown to teachers and testers, many assigned topics prove inaccessible to many student writers. Such topics may cause students to appear disabled or to be among the "least able students" when they are in fact not. Ways are needed to determine the accessibility level of topics and to create accessible topics.

Empirical evidence of the effects of topics on writers also exists. Freedman (1981) and Freedman and Calfee (in press) have shown that in large-scale testing situations, topics intended to be parallel in difficulty and accessibility significantly influenced the scores received by college-age test takers. Keech (1982) reports how topics used in different years for a writing proficiency examination at the high school level elicited different types of writing from the testees. Keech found that the topics that were designed to be parallel in difficulty in fact were not parallel. The only topic that is justifiably inaccessible is one that is at a particular level of difficulty and because of that difficulty proves inaccessible to some. Once testers learn how to keep a topic's accessibility level constant while manipulating its difficulty level, it will be possible to create topics that can accurately discriminate between students at different levels of writing proficiency. As a first step toward achieving this goal, Ruth (1982) is developing techniques for analyzing topics to determine whether or not they are parallel.

Recommendations

Because of the difficulty in getting samples of writing that are comparable across occasions, Braddock, Lloyd-Jones, and Schoer (1963) and Diederich (1974) recommend eliciting several writing samples before judging students' writing levels. This advice seems sound considering the potential effects on a student's performance of (a) the classroom and testing contexts, (b) the unpredictable variations in a writer's mood and attitude toward writing during the writing process, and (c) the writing topic. The more samples of writing one collects and the more varied the conditions under which they are collected, the better the chance of accurately determining how capable the student is. We must take great care when testing young children to find out whether they have "writing disorders," and if they do to decide on a program for helping children identified as having such disorders.

COMPARABILITY OF JUDGMENTS AND CHOICE OF SCORING METHOD

Holistic

When students produce pieces of writing, they must be evaluated. Obviously, a story or an essay cannot be rated as right or wrong, but someone must judge it. In the history of writing assessment, all approaches to scoring pieces of writing have confronted the problems of getting

evaluators to agree on their judgments of a single piece of student prose. Cooper (1977) outlines the holistic method of judging test essays used by the Educational Testing Service (ETS) and the College Entrance Examination Board (CEEB), in which an essay is given a numerical score that represents its *overall* quality. To get evaluators to agree on these scores, ETS trains them on the standards set by ETS-chosen leaders of the rating (a room leader and a group of table leaders). During training, the leaders present the raters with preselected essays that represent the range of quality in the given group of essays. These essays are called *anchor papers* and are meant to be prototypes for each score on the scale. Since the standards are set by referring only to the existing set of papers, one must be able to assume that a normal population is being tested on each testing occasion, and that the standards can be determined from the given set of papers. If the population is always similarly distributed, the scores from one testing occasion to the next should be comparable.

The leaders are responsible for keeping the raters in agreement with the standards established by the anchor papers. Using this rigorous procedure, ETS and CEEB have managed to achieve high levels of interrater reliability. Such a procedure, though, requires rigorous training of raters and the assumption of a normally curved group of papers from which to draw samples for anchor papers. Without such training, even a homogeneous group of raters will not naturally agree with one another (Diederich, French, & Carlton, 1961).

Analytic

Diederich and his colleagues determined that the difficulties in achieving interrater reliability had their source in the values that different raters ascribed to various features of the text. To overcome this problem, Diederich (1974) developed an analytic rating scale in which each part of the text is rated separately and then the different scores for each category of the scale are weighted. His categories include ideas, organization, wording, flavor, usage and sentence structure, punctuation, capitalization, abbreviations, numbers, spelling, handwriting, and neatness.

This analytic method of judging essays seems appealing, especially for those interested in diagnosing specific writing problems; however, for several reasons, it has not had much popularity among testers. First of all, there is no evidence that raters will agree any better on these subscores of the analytic scale than they will on holistic score. Hirsch (1977) argues that Diederich's (1974) system only codifies disagreement. Also, the method is time-consuming, and the subscores generally correlate highly with one another, meaning that they yield little information beyond what the single holistic score offers (Freedman, 1981). Nevertheless, Diederich describes each category clearly and gives good directions for how to score a paper on each category. His category descriptions have proven helpful to teachers of writing as they develop standards for judging the writing of their students and to students who want to understand the standards teachers apply when they judge essays.

Primary trait

Another approach to rating essays comes from the National Assessment of Educational Progress (NAEP), which developed a method of primary trait scoring (Lloyd-Jones, 1977). This method yields reliable results, but the judgments do not necessarily relate to the quality of the piece of writing. In the primary trait method, the tester decides ahead of time what type of response is appropriate on a given writing task. The primary traits consist of statements of what the writer has to do to fulfill the demands of the task. For example, if the task requires the writer to compose a business letter, one of the primary traits might be the proper opening and closing form for a business letter. Only those who used the proper form would receive credit for that trait. The primary trait method of scoring has the advantage of offering information beyond a single score denoting overall quality and of providing scores on which raters can agree; however, those interested in diagnosing and treating writing disorders will need more detailed and different information from that which primary trait scores offer.

Postholistic feature analysis

Keech and Blickhahn (1982) offer suggestions for analyzing the features of essays after they have been scored holistically. They emphasize that their technique is useful only as a post hoc analysis of the holistic score, not as a scoring technique. They offer procedures for analyzing and diagnosing essays, procedures that could be replicated to fit individual testing settings. They describe one such feature analysis developed to explain the difference between borderline passing and failing papers on a high school competency test; the critical features distinguishing the passing from failing papers included point of view, development and concreteness level, coherence and paragraphing, sentence sense, word choice, punctuation and capitalization, spelling, legibility and presentation, and word production. Once features have been identified, papers are rescored for adequacy in the given features. It is important to note that Keech and Blickhahn emphasize that the features are selected according to the dictates of a particular situation. They claim that their system is different from Diederich's (1974) analytic scale in that it is designed to give formative feedback to students rather than to provide a summative evaluation.

Recommendations

When choosing a scoring method, testers must decide on the purpose for the evaluation. Spandel and Stiggens (1980) discuss the following purposes of writing evaluation: instructional management—diagnosis of learner strengths and weaknesses, placement, guidance, and counseling; program management—admission to particular classes or programs and proficiency certification; program evalua-

When choosing a scoring method, testers must decide on the purpose of the evaluation.

tion—survey assessments, formative evaluation, and summative evaluation. Like Keech and Blickhahn, they emphasize that the context will dictate the scoring method.

VALIDITY OF INDIRECT MEASURES

Largely in response to the difficulties of getting raters to agree with one another and to the time-consuming and expensive nature of getting people to evaluate writing samples, some have attempted to devise ways to test writing ability indirectly. In other words, testers have attempted to develop multiple-choice, machine scorable tests of writing. Obviously, such tests, if they measure writing ability at all, measure it indirectly, since it is impossible to gather a writing sample in a multiple-choice format. This format, testers argue, offers reliability because students have a chance to respond to multiple items rather than to one or two essay items. Although this argument has merit, it is still difficult to justify testing writing without gathering any writing sample. Teachers have opposed by movements toward indirect assessment, and recently the large scale testers such as ETS, CEEB, NAEP, and many statewide proficiency and placement testing programs have heeded the teachers' opposition by including writing samples alongside multiple-choice questions. Nevertheless, essays need to be collected in different settings and on different topics that call for different types of writing to assess writing proficiency adequately, a task beyond the capacity of the large-scale tester but not of those who test a few children for special purposes. Instead of focusing a great deal of energy on scoring a single writing sample, as Myklebust (1965) and Poteet (1980) recommend, it is probably wiser first to examine several writing samples in less detail.

WRITING PROBLEMS AND REMEDIES

Once a writing problem has been identified, or during the process of identifying a problem, it is important next to discover the specific nature of the problem and the reasons for its existence. Only through an analysis of these reasons can the most appropriate remedies be found. In Shaughnessy's (1977) book about the writing problems of students admitted to the City University of New York in its open admissions program during the 1960s, she discusses the sources of these students' dire problems. The discussion provides an excellent general framework for discovering the sources of student writing problems.

Before presenting Shaughnessy's framework, some background information about her students' writing would be helpful. She calls these college-age students basic writers (BW); their writing problems are blatant. Some BWs have not made the transition from printing to cursive; some struggle with the motor coordination required to put letters and words onto paper; many have severe difficulties with spelling, punctuation, capitalization, written syntactic patterns, and written discourse level patterns.

Shaughnessy stresses a point that is critical to all special educators involved in the diagnosis and remediation of children

with serious writing problems: The teacher's (and the tester's) most important job is to discover "the logic" underlying the student's problems. She shows how logical the problems of her BW students are by providing a framework for discovering the sources of their syntactic problems. She finds three different sources: (a) the student may not have internalized the "*language patterns* characteristic of written English" (p. 73); (b) the student may be unfamiliar "with the *composing process*" (p. 73); and (c) the student may have difficulties with his or her "*attitude* toward himself within an academic setting" (p. 73). Just as this framework can be applied to all writers with problems, it also applies beyond the level of syntax. Shaughnessy (1977) continues by making recommendations about teaching:

> Each of these explanations [for the student problem] suggests a pedagogy: The pedagogy that stresses grammar, whether in the abstract or as a set of forms to be generated through practice with sentences, tends to assume that students do not have command of many of the forms required in written English and must therefore learn them through explicit instruction; the pedagogy that stresses process (prewriting, free writing, composing, re-scanning, proofreading, etc.) tends to minimize the value of grammatical and rhetorical study and assume, rather, that students already "know" the wanted forms but cannot produce them, nor anything resembling their own "voices," until they are encouraged to *behave* as writers; the pedagogy that stresses the therapeutic value of writing and seeks the affective response to whatever is read or discussed tends to see *confidence* as central to the writing act and to dismiss concerns with form or process as incidental to the students' discovery of themselves as individuals with ideas, points of view, and memories that are worth writing about. A teacher should not have to choose from among these pedagogies, for each addresses but one part of the problem. (p. 73)

Ideally, those involved in identifying and remedying writing disorders will develop instruments for the assessment of writing that will fulfill several requirements: the instruments will identify true cases of disorder, the instruments will specify what the disorder is, and this specification of the disorder will imply a pedagogical program. Unfortunately, I know of no commercially available tests of writing that fulfill these requirements. For now we must rely on the judgments of expert teachers and testers who are capable of providing children with opportunities to communicate in writing, supporting children throughout the writing process, evaluating their performances, and teaching them to improve.

REFERENCES

Applebee, A. *Writing in the secondary school: English and the content areas*. Urbana, Ill.: National Council of Teachers of English, 1981.

Braddock, R., Lloyd-Jones, R., & Schoer, L. *Research in written composition*. Urbana, Ill.: National Council of Teachers of English, 1963.

Cazden, C., John, V., & Hymes, D. *Functions of language in the classroom*. New York: Teacher's College Press, 1972.

Cooper, C. Holistic evaluation of writing. In C. Cooper & L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging*. Urbana, Ill.: National Council of Teachers of English, 1977.

Diederich, P. *Measuring growth in English*. Urbana, Ill.: National Council of Teachers of English, 1974.

Diederich, P., French, S., & Carlton, S. *Factors in judg-*

ments of writing ability (Research Bulletin RB-61-15). Princeton, N.J.: Educational Testing Service, 1961.

Dyson, A. Viewpoints. *Language Arts*, 1980, 57, 725–726.

Dyson, A. Oral langauge: The rooting system for learning to write. *Language Arts*, 1981, 58, 776–784.

Freedman, S. How characteristics of student essays influence teachers' evaluations. *Journal of Educational Psychology*, 1979, *71*, 328–338.

Freedman, S. Influences on the evaluators of expository essays: Beyond the text. *Research in the Teaching of English*, 1981, *15*, 245–255.

Freedman, S., & Calfee, R. Holistic assessment of writing: Experimental design and cognitive theory. In P. Mosenthal, L. Tamor, & S. Walmsley (Eds.), *Research in writing: Principles and methods*. New York: Longman, in press.

Graves, D. *Balancing the basics: Let them write*. New York: Ford Foundation, 1978.

Grobe, S. Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings. *Research in the Teaching of English*, 1981, *15*, 75–85.

Hirsch, E.D. *The philosophy of composition*. Chicago: University of Chicago Press, 1977.

Howerton, M., Jacobson, M., & Selden, R. *The relationship between quantitative and qualitative measures of writing skills*. Charlottesville: University of Virginia, 1977 (ERIC Document Reproduction Service No. ED 137 416).

Keech, C. *Unexpected directions of change in student writing performance*. Paper presented at the annual meeting of the American Educational Research Association, New York, March 1982.

Keech, C., & Blickhahn, K. *National writing project guide to holistic assessment of writing*. Berkeley: Bay Area Writing Project, University of California, 1982.

Lerner, J. *Children with learning disabilities* (3rd ed.). Boston: Houghton Mifflin, 1981.

Lloyd-Jones, R. Primary trait scoring. In C. Cooper & L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging*. Urbana, Ill.: National Council of Teachers of English, 1977.

Loban, W. *Language development: Kindergarten through grade twelve* (Research report No. 18). Urbana, Ill.: National Council of Teachers of English, 1976.

Moran, M. Performance of learning disabled and low achieving secondary students on formal features of a paragraph-writing task. *Learning Disability Quarterly*, 1981, *4*, 271–280.

Myklebust, H. *Development and disorders of written language 1*. New York: Grune & Stratton, 1965.

Nold, E., & Freedman, S. An analysis of readers' responses to student writing. *Research in the Teaching of English*, 1977, *11*, 164–174.

Poplin, M., Gray, R., Larsen, S., Banikowski, A., & Mehring, T. A comparison of components of written expression abilities in learning disabled and non-learning disabled students at three grade levels. *Learning Disability Quarterly*, 1980, 3(4), 46–53.

Poteet, J. Informal assessment of written expression. *Learning Disability Quarterly*, 1980, 3(4), 88–98.

Ruth, L. Properties of writing tasks: *Student and scorer interpretations in holistic writing assessment events* (Final Rep. NIE G-80-0034, "A study of procedures for conducting holistic writing assessment," James Gray, principal investigator) Berkeley: Bay Area Writing Project, University of California, June 1982.

Shaughnessy, M. *Errors and expectations: A guide for the teacher of basic writing*. New York: Oxford University Press, 1977.

Slotnick, H., & Knapp, J. Essay grading by computer: A laboratory phenomenon? *English Journal*, 1971, *60*, 75–87.

Spandel, V., & Stiggins, R. *Direct measures of writing ability: Issues and applications*. Portland, Oreg.: Northwest Regional Educational Laboratory 1980.

Wells, C.G. *Learning through interactions: The study of language development*. Cambridge, England: Cambridge University Press, 1981.