Educational Psychology Review, Vol. 6, No. 2, 1994

A Cognitive Theoretical Approach to Reading Diagnostics

James M. Royer[1,3] and Gale M. Sinatra[2]

After a brief review of current reading diagnostic practices, the article presents a cognitive developmental theory and indicates how the theory can be used to identify potential sources of reading problems. Four criteria for a reading diagnostic system based on cognitive developmental theory are then identified: (1) the system must provide reliable and valid assessments, (2) performance on the diagnostic system must be consistent with cognitive developmental theory, (3) the diagnostic procedures must provide specific information about the nature of the reading difficulty the student is experiencing, and (4) diagnoses provided by the procedures must lead to prescriptive procedures that alleviate to a demonstrable degree the reading problems the student is experiencing. The Computer-based Academic Assessment System (CAAS) is then introduced, and evidence from three studies is used to examine the extent to which the CAAS system can satisfy the four requirements for a theory-based reading diagnostic system. The article closes with a discussion of the empirical evidence and with suggestions for future research.

KEY WORDS: reading diagnostics; dyslexia; reading development; cognitive development.

INTRODUCTION

The majority of the students enrolled in our schools learn to read without major difficulty. However, approximately 20–25% of students do experience moderate to severe difficulties in learning to read (Stedman and Kaestle, 1987). Students with reading problems consume an enormous pro-

[1]University of Massachusetts, Amherst.
[2]University of Utah.
[3]Correspondence should be directed to James M. Royer, Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01003.

1040-726X/94/0600-0081$07.00/0 © 1994 Plenum Publishing Corporation

portion of societal resources during their school years and after they leave school. While in school, children with reading problems are targeted by federal programs (e.g., Chapter 1 of the Elementary and Secondary Education Act), by State initiatives, and by local special education programs. After children with reading problems leave school they constitute a large proportion of the unemployed and disadvantaged. While the cause of some children's reading difficulties may be organic in nature rather than experiential, there is currently a move toward viewing reading difficulties on a continuum from "garden-variety poor readers" (Gough and Tunmer, 1986) to true dyslexics (see Stanovich, 1988 for the distinction). Given this orientation, early and accurate diagnosis of reading difficulties is crucial. Left unrecognized, poor readers can acquire "increasingly global cognitive deficits" as they get older (Stanovich, 1988).

Given the dire consequences of reading problems, it is no surprise that an enormous amount of activity is devoted to the diagnosis and remediation of reading difficulties. What is surprising though, given both the wealth of new knowledge evident in the literature and the number of diagnostic tests available, is that many of the diagnostic and remedial procedures currently in use are generally ineffective (Vinsonhaler, Weinshank, Wagner, and Polin, 1983). Vinsonhaler *et al.* (1983) argue that the techniques are highly unreliable in their diagnosis of problems and tend to be uninformative when it comes to prescriptives for remediation. They dramatically illustrated this point in a series of six studies. Vinsonhaler *et al.* (1983) demonstrated that the agreement of two reading diagnosticians when examining the same reader was only .1, and when the same diagnostician examined the same reader twice (under different names) the agreement was only .2. Moreover, there was no consistency of recommendations for remediation given the same diagnostic decision. Others have noted the lack of correspondence between diagnostic decisions and remedial prescriptions (e.g., Aaron, 1991).

One likely reason for the inconsistency of diagnostics and remedial recommendations is the lack of relationship among theories of reading, diagnostic assessment, and reading instruction and remediation. Rather than theories of reading informing the development of diagnostic instruments and instructional techniques, which then suggest remedial approaches, it is generally the case that research in each of these areas is conducted independent of the others. Even in cases where diagnostic tests are based on a theory of reading, the relationship between the theory and the nature of the diagnostic procedure is tenuous at best. Given the lack of connection between reading theory, diagnosis, and remediation, reading specialists are left to their own devices when it comes to developing instructional approaches designed to remediate problems.

The lack of correspondence between theory and practice has led to a call for the development of theory-based assessments that will have practical implications for diagnosis and remediation (e.g., Glaser, 1981; Sternberg, 1981; Schwartz, 1980; Brown and Campione, 1986; Linn, 1986; Glass, 1986). In the next section of the paper we will briefly examine current diagnostic procedures. We will then describe a theory of reading that has direct implications for reading diagnostics and remediation. This discussion will be followed by the description of a diagnostic system motivated by the theory, and by a presentation of data that evaluates the diagnostic potential of the system.

TRADITIONAL DIAGNOSTIC ASSESSMENT OF READING

Reading is a complex process and the difficulties inherent in diagnosing reading problems reflect that complexity. Generally, when a child is having difficulty reading, the first step is to determine if there are any physical disabilities such as visual or auditory deficiencies, or neurological disorders. After ruling out these potential causes; emotional, social, and environmental factors must be considered. If these factors do not appear to be the problem, typically the next step is to administer some type of standardized test of general intelligence.

A child is not considered reading disabled or dyslexic if he or she is found to be low in general intellectual functioning; rather, a child is considered reading disabled if there is a discrepancy between his or her expected reading capability and his or her actual reading performance (Bejar, 1984; Stanovich, 1988).

There is increasing criticism of this approach to reading diagnostics (Stanovich, 1988). First, the use of intelligence tests in making diagnostic decisions has come under considerable attack (Aaron, 1991). Second, the determination of a child's "expected" performance is a process having questionable reliability and validity. Third, the determination of a child's actual reading level is often assessed with standardized reading tests. There has been growing criticism of using standardized tests to assess reading ability (Johnston, 1992). Further, standardized reading tests are particularly problematic when used for diagnostic purposes. Although they may appear to have diagnostic utility (they often contain subtests of word decoding and comprehension) in actuality, the tests yield little diagnostic information. The basic problem stems from the psychometric properties of standardized reading tests; that is, they are designed to compare an examinee's performance to norms, rather than providing information about the reading capabilities of a particular reader.

Although intelligence tests and standardized reading tests are usually used to determine whether a child has a reading problem, typically, an actual diagnosis of the problem will involve the administration of an informal diagnostic test (see, for example, The Reading Miscue Inventory, 1972) or a formal diagnostic test (see, for example, the Durrell Analysis of Reading Difficulties, 1980). Diagnostic tests are typically based on an analysis of a child's oral reading capability.

Oral diagnostic tests involve having the child read aloud text segments of varying difficulty. The examiner records the errors the reader makes during this process and then attempts to place the errors into categories. The examiner then consults the test manual for a diagnosis based on the nature and the frequency of the errors the reader has made. The diagnostic decisions may be useful when the nature of the child's difficulty corresponds to a diagnostic category. However, many difficulties that readers have may not be revealed by oral reading performance (see Rayner and Pollatsek, 1989, p. 181 for a discussion of the difficulties inherent in inferring reading processes from oral reading errors). Further, many children have accurate decoding skills but do not understand what they read. Their reading difficulties would not be revealed through oral reading errors. These children may be accurate decoders but have slow lexical-access abilities, prior knowledge or metacognitive problems, or other problems which would not be revealed by examining the accuracy of their performance on oral reading tests.

A COGNITIVE DEVELOPMENTAL THEORY OF READING

The general outline of the theory that will be described in this section has been evolving for a number of years, and a number of researchers have made contributions to the theory (e.g., Anderson, 1983; Fodor, 1983; Forster, 1979; Perfetti, 1988; 1992, Royer, 1985; Stanovich, 1986a; 1990; van Dijk and Kintsch, 1983). The particular version of the theory described here borrows elements from these theorists and others.

Enabling Skills

Cognitive Developmental Theory suggests that successful reading is dependent on at least two enabling skills: phonological awareness and the ability to identify letters (e.g., Stanovich, 1986a; Bryant and Bradley, 1985; Wolf, 1984). Phonological awareness refers to the recognition that speech sounds can be subdivided into constituent elements called phonemes. An example of a task that assesses phonological awareness would be to orally present a child with a word (e.g., tall) and then ask the child to say which of the following words (lark, hall, milk) has an ending that sounds like the end of the first word presented. There is an enormous literature indicating that the ability to accomplish this task is predictive of future reading success and that pre-reading children instructed in accomplishing the task show reading benefits relative to children who have not been instructed (see Stanovich, 1986a; Brant and Bradley, 1985 for reviews).

The second reading enabling skill is letter recognition. As is the case with phonological awareness, there is a very large literature indicating that letter recognition is a good predictor of subsequent reading performance in pre-reading children (Stanovich, 1986a; Wolf, 1984). Phonological awareness and letter recognition are called enabling skills because the theory suggests that both phonemic awareness and letter perception are essential foundations for subsequent reading development.

Word Identification

Having become aware of phonemic elements in speech and having developed the ability to accurately recognize letters, children then begin to develop the capability of mapping phonemes onto letters and letter combinations (or spelling patterns). This ability forms the foundation for the acquisition of word-identification skills.

The development of word identification skills occurs within a cognitive system that is constrained by a relatively small working-memory capacity. Working memory can be thought of as that part of memory that constitutes conscious awareness. Theorists vary in their estimates of the actual size of working memory, but most would place the average limits somewhere between four and eight elements. Moreover, elements will decay from working memory within a relatively short period of time (about 10–15 seconds) unless they are maintained by conscious rehearsal.

Early stages of the development of word identification are conducted as if the child were solving a problem. The child may use word envelope forms, surrounding sentence or pictorial context, prior memory of stories, or strategic guessing to identify printed words (Perfetti, 1992; Gough, Juel, and Griffith, 1992). The process is relatively slow and the child uses a large part of his/her working-memory capacity to decipher words. At this stage of reading development, word recognition is very much a "top-down" process. That is, the child uses a number of conscious strategies and considerable prior knowledge in recognizing a word (Perfetti, 1992).

As exposure increases, recognition of frequently encountered words increasingly becomes a "bottom-up" or "data-driven" process. The reader no longer has to think about what a word is; the mere sight of the word on the page is sufficient to trigger recognition of the word (see Ehri, 1991; 1992 for a review of the development of word-recognition ability). There are several consequences associated with this stage of development. Word identification becomes very fast, essentially load free, and impervious to higher level cognitive activities (Perfetti, 1992; Stanovich, 1990). Each one of these developments contains important implications for the reading process.

Rapid processing of words means that the reader can hold more words in working memory within a limited time period, thereby increasing the probability that a meaningful chunk (e.g., idea unit, phrase) will be input before initially processed words decay from memory. Load-free processing means that cognitive capacity which initially was tied up with trying to decipher letter strings can now be devoted to higher level comprehension and thought processes. And finally, identification of some words becomes an "encapsulated" (Fodor, 1983; Forster, 1979) activity that cannot be affected by context or by conscious, strategic, activities.

Encapsulation of processing does not occur with all words. Children who are developing readers will exhibit encapsulated processing of some words and conscious processing of others. And even highly skilled readers will occasionally encounter words that require a switch from data-driven mode to conscious-attention mode. This commonly occurs with unfamiliar words or with words that are stimulus degraded (Rayner and Pollatsek, 1989). Stimulus degraded words are those where the type is dim, letters are cut off or missing, or some other aspect of the type of print makes it difficult to read the words.

Activation of Meaning

An important aspect of skilled reading is that meanings of words are automatically activated immediately after the word is recognized. Swinney (1979) demonstrated that multiple common meanings of a word are activated even though the context in which the word appears supports only one of the meanings. As an instance, if a skilled reader were to read, "the bug walked down the wall," both bug as listening device and bug as insect meanings would be activated. Selection of the appropriate meaning occurs only after multiple meanings of a word have been activated. The speed of access to semantic memory has proven to be an important discriminator between skilled and less-skilled readers (Hunt, Lunneborg, and Lewis, 1975; Chabot, Petros, and McCord, 1983).

Syntactic and Semantic Processing

Most theorists view the syntactic processing of spoken language as an encapsulated activity that is "impenetrable" by other cognitive activities (e.g., Fodor, 1983; Forster, 1979; Perfetti, 1992). Some theorists take the added step of suggesting that the cognitive component, or "module" as they refer to it, is part of our genetic endowment as human beings (e.g., Fodor, 1983). Syntactic processing ability is generally fully developed through speech experience in children free from language impairment by the time they begin to develop reading skills, and therefore is unlikely to be a source of reading difficulty.

Semantic processing, or text modelling as it is sometimes called, requires the simultaneous utilization of a number of higher level cognitive capabilities (e.g., Perfetti, 1988; van Dijk and Kintsch, 1983). In order to interpret text information, readers must be able to relate printed information to prior knowledge, they must be able to make automatic (unconscious) and conscious inferences required to knit the details of a text together, and they must be able to consciously connect information being read to related information in memory. In addition, they must be aware of whether or not they are understanding what is being read, they must be able to change their reading behavior if they are not understanding, and they must be able to use a variety of reading strategies depending on the nature of the text being read and the purpose for reading.

Prior Knowledge and Metacognitive and Strategic Processing

Cognitive developmental theory suggests that reading comprehension is a constructive process (Spiro, 1980). The constructivist perspective asserts that meaning is not inherent in a text. Instead, meaning must be constructed in a process involving the interaction of the message contained in a text, the prior knowledge of the reader, and the context (both textual context and environmental context) in which the reading act occurs. Sometimes readers fail to understand, not because they have poor reading skills, but rather, because they do not have the prior knowledge necessary to interpret the text (Anderson and Pearson, 1984; Anderson, Reynolds, Shallert, and Goetz, 1977).

Recently, the importance of metacognitive and strategic processing in reading comprehension has been emphasized (Baker and Brown, 1984;

Pressley, Goodchild, Fleet, Zajchowski, and Evans, 1989; Wade, Schraw, Buxton, and Hayes, 1993). It seems that having adequate processing skills is insufficient if readers fail to monitor their comprehension process and/or fail to use their processing resources efficiently (Wade *et al.*, 1993).

Implications of Cognitive Developmental Theory for Reading Diagnostics

The theory sketched in the preceding section suggests a number of reasons for the development of reading problems. These reasons can be grouped into three major categories: failure to develop critical component processing skills that lead to the development of skilled reading, inadequate prior knowledge, and failure to develop and/or use efficient metacognitive and strategic processes.

Failure to Develop Critical Component Process Skills. Children may develop difficulties in learning to read because of inadequate development of a critical component process such as: phonological awareness, letter identification, or word-recognition ability. Inadequate development of these skills could have a profound effect on subsequent reading development.

Several researchers (e.g., Walberg and Tsai, 1983; Stanovich, 1986a) have written of the cascading of positive and negative effects that accompany the adequate or inadequate development of enabling and word-recognition skills. The existence of adequate enabling and word-recognition skills allows a developing reader to begin to acquire a vocabulary and to relate that vocabulary to conceptual knowledge acquired via experience. The acquisition of a vocabulary, and the development of an ever increasing ability to decipher words and their meaning, increases the size of the reading vocabulary and the number of entries in conceptual memory. These increases in competence are mediated by an enormous amount of practice. One researcher estimates that a skilled reader (one who has presumably mastered the phonological awareness and letter enabling skills) in the first grade reads 1933 words *a week* (Allington, 1984). By the middle grades, an average reader may be reading 1 million words per year and the voracious, highly skilled, reader might be reading as many as 50 million words a year (Nagy and Anderson, 1984). In contrast, first grade readers with poor reading skills may read as few as 16 words a week (less by a factor of over 100 compared to skilled readers), and poorly prepared or motivated readers in the middle grades may read 100 times fewer words than more capable readers.

These differences produce what have been called, "Matthew effects" (Walberg and Tsai, 1983; Stanovich, 1986a). Named after the biblical passage, Matthew XXV:29, Matthew effects describe a situation where students with adequate word identification skills and students with poorly developed word recognition skills develop enormous differences in educational competence over a relatively short period of time. Students with adequate enabling and word recognition skills quickly acquire the rudiments of reading ability and this in turn accelerates the acquisition of further competence. In comparison, students who do not possess adequate enabling skills begin with a disadvantage that inhibits the development of word identification skills; poor word identification skills retard vocabulary development; and inadequate vocabulary development retards knowledge acquisition, and so on. The net result is an enormous gap in educational competence between students initially having enabling skills and those who do not have enabling skills.

The cognitive developmental theory of reading suggests that readers must not only learn to identify words but must develop the ability to identify words very rapidly and accurately. Moreover, the process of word identification must occur with minimal demands on cognitive capacity. This suggests that one source of reading difficulty would be the failure to develop automaticity in the word-identification process. The failure to develop automatic processing of frequently encountered words would affect reading performance in two ways: (1) because words can only exist for a limited time in working memory, slow word identification would result in some words decaying before a meaningful chunk is processed, and (2) conscious word identification activities consume valuable cognitive capacity, capacity that could otherwise be directed toward higher-level comprehension activities.

The relationship between enabling skills and reading development suggests that a check for the acquisition of enabling skills should be a routine diagnostic procedure for young children who are suspected of being at-risk for the development of reading problems. Early detection of problems accompanied by instructional procedures designed to develop enabling skills can result in enhanced reading capability for the average pre-school student (Ball and Blachman, 1988) and for children who have been identified as being at risk (Bradley and Bryant, 1983; Bradley, 1988).

The preceding analysis also suggests that another diagnostic activity that might yield valuable information regarding the source of reading problems would be to examine the extent to which readers could identify words rapidly, accurately, and with minimal demands on cognitive capacity. A reader who exhibited deficiencies in word processing could then engage in instructional activities designed to increase speed and accuracy of performance. For instance, Frederiksen, Warren, and Rosebery (1985a, b) have reported research showing that computer-based gaming activities designed to increase speed and accuracy of word processing produced a number of beneficial effects in junior-high school students with reading problems.

Inadequate Prior Knowledge. Another potential source of reading problems is inadequate prior knowledge. Comprehension difficulties due to inadequate prior knowledge or prior knowledge that conflicts with the content of the text have been demonstrated extensively (Anderson *et al.*, 1977; Posner, Strike, Hewson, and Gertoz, 1982). Royer and Cunningham (1981) have described situations involving children from impoverished backgrounds where it is impossible to evaluate the meaning of poor performance on a standardized reading comprehension test. The poor performance could mean that the student has inadequate skills, but it could also mean that the student did not have the background knowledge necessary to construct an interpretation of the text. Without additional information, either of these interpretations is possible. Similar problems have been pointed out in evaluating the standardized test performance of children from different cultural and linguistic backgrounds (Mestre and Royer, 1991; Royer and Carlo, 1991). Because standardized reading tests are often used for diagnostic purposes, the issue of prior knowledge is a critical one for diagnosis of reading difficulties.

Metacognitive and Strategic Processing Inadequacies. Some children appear to develop reading comprehension difficulties because of poor metacognitive abilities. Children with poor metacognitive abilities exhibit problems that can be grouped into three clusters (Baker and Brown, 1984). The first cluster consists of personal knowledge about the availability and use of cognitive resources as they relate to a particularly reading activity. Awareness, for instance, that organized material is easier to learn than unorganized material and that one is capable of applying different learning approaches to the two types of material represents an instance of metacognitive knowledge.

The second cluster of metacognitive capabilities involve self-regulatory activities a learner engages in while problem solving. Baker and Brown (1984) suggest that active learners *check* the outcome of a learning act, they *plan* their next activity, and they *monitor* the effectiveness of any strategy change. Furthermore, they point out that the failure to constantly monitor the outcome of the reading and learning process have serious consequences with respect to effective learning.

The final cluster of metacognitive activities that Baker and Brown (1984) describe relates to the development and use of compensatory reading strategies. Good readers must develop the ability to read technical material that must be learned in detail differently than they would read light fiction. Moreover, they must immediately recognize that one strategy is called for in one situation and a different strategy in another.

A number of researchers have shown that students instructed in strategic reading activities subsequently improve in reading performance (e.g., Paris, Cross, and Lipson, 1984; Palincsar and Brown, 1984). The direct implication of these studies is that it should be possible to identify children with strategic reading difficulties, to identify the nature of those difficulties, and to provide instructional interventions that remediate the difficulties, thereby improving reading performance.

The Repair of a Lower Level Skill or a Higher Level Skill Does Not Necessarily Result in Improvements in Overall Reading Performance. The discussion in the preceding sections suggested that a complete diagnostic system would include assessments of enabling skills, word identification skills, metacognitive abilities, knowledge and utilization of reading strategies, and availability of the prior knowledge necessary to interpret texts of interest. Given poor performance in any of these assessment areas, the research literature suggests that educational interventions can attack the problem with resulting benefit to reading performance. It should be noted, though, that the repair of one inadequate cognitive skill may not necessarily lead to improvements in overall reading performance. Rather, the repair of a lower level block may serve to remove an impediment that prevents the development of upper level capabilities, but those upper level capabilities must be developed after the lower level impediment has been removed. Likewise, the training of higher level activities like the skilled use of metacognitive strategies may not enhance overall reading performance if students have problems at a lower level.

This analysis suggests that the diagnostic and instructional process should be continuous for readers having difficulties. That is, the lowest level deficit should be identified and repaired, followed by a reevaluation of the reader for additional problems, and by further instructional intervention to repair newly identified problems. It should be noted that the repair of lower level skills can be accomplished using a variety of instructional methods. As an instance, suppose that a diagnostic procedure indicated that a reader was very deficient in speed of word identification. One way to repair the deficient skill would be to provide the child with intensive practice on the identification of isolated words. But the same end could probably be accomplished by increasing the amount of reading activity the child engages in. Increased reading practice increases the efficiency of reading processes; hence, increased activity directed at *reading comprehension* could also result in improved word identification performance. Another instructional activity might be directed at teaching the child strategies like paying attention to theme and context as a means of assisting in the word identification process.

The important point to be made is that the repair of deficient skills can be accomplished using a variety of instructional techniques. The technique being used, though, should be evaluated by constantly checking the status of the deficient skill, and should be changed if there is inadequate progress in repair of the skill. It is probable that instructional techniques will interact with the nature of the difficulty being experienced by the reader, and this will be an important research area in years to come.

The idea that diagnosis and instructional intervention may need to be continuous along the entire range of reading skills may help to explain why some efforts designed to enhance reading performance appear to have greater success than others. For instance, several efforts involving the identification of young children who are deficient in phonological awareness skills followed by instructional intervention in the missing skills (e.g., Bradley and Bryant, 1985; Lundberg, Frost, and Peterson, 1988) have shown that instructed children are subsequently better readers (when reading skills develop) than are uninstructed readers. It should be noted that these studies involve identifying poorly developed enabling skills that may block the subsequent development of skilled reading performance, repairing those skills, and then letting the normal course of reading development occur.

In contrast to the success stories involving very young children, other studies directed at repairing skills in older children have been less successful. For instance, the Frederiksen *et al.* studies (1985a, b) conducted with poor reading junior-high school students resulted in significant improvement in the processing speed of lower level activities (e.g., multi-letter processing, phonological decoding, use of context in retrieving and integrating word meanings), but had a comparatively smaller impact on measures involving the reading and understanding of sentence length material. Likewise, Roth and Beck (1987), who sought to improve the word-decoding speed of poor fourth grade readers found that significant improvements in word-decoding speed enhanced word and sentence comprehension but did not improve comprehension at the paragraph level. Finally, the Paris *et al.* (1984) study that trained third and fifth grade children in the use of metacognitive reading strategies showed that improved strategy usage did not necessarily result in improvements on standardized tests of reading comprehension.

The studies described above that do not show an overall improvement in reading performance involve focusing on reading skill, repairing that skill, and then looking for an immediate impact on overall reading performance. This research strategy may miss improvements that occur some time after the instructional intervention has occurred. In addition, the studies do not evaluate the potential positive impact that might occur in a situation where there was a diagnostic identification of the lowest flawed reading skill, the subsequent repair of that skill, followed by an evaluation of the next skill in line, the repair of that skill if necessary, and so on. Again, it should be pointed out that this strategy does not dictate the particular instructional approach chosen to repair the skill. Rather, it suggests that deficient skills should be identified and that the success of the instructional approach taken should be evaluated in terms of its success in repairing the deficient skill.

SUCCESS CRITERIA FOR A DIAGNOSTIC PROCEDURE BASED ON COGNITIVE/DEVELOPMENTAL THEORY

The review of cognitive developmental theory and the consideration of the diagnostic and instructional implications of the theory suggests a framework for the development of systems that have diagnostic and prescriptive value. A remaining question involves determining how one would recognize a system that fulfilled these goals. We suggest there are four criteria that diagnostic procedures based on cognitive developmental theory must meet: (1) the procedures must be reliable and valid, (2) the procedures must yield patterns of performance that are consistent with cognitive developmental theory, (3) the procedures must provide specific information about the nature of the reading difficulty the student is experiencing, and (4) diagnoses provided by the procedures must lead to prescriptive procedures that alleviate to a demonstrable degree the reading problems the student is experiencing.

It should be apparent that no single diagnostic procedure will accomplish the full range of necessary diagnostic activities. Moreover, it should also be apparent that no signal study will be likely to satisfy all four of the criteria mentioned above. We have made an attempt, however, to initiate research in this area. In the following section we describe an assessment system that is motivated by cognitive developmental theory and we present some preliminary results that are relevant to evaluating the diagnostic adequacy of the system. Following presentation of the results, we evaluate those results in terms of the four criteria for a successful diagnostic system.

A COGNITIVE THEORETIC APPROACH TO READING DIAGNOSTICS: THE COMPUTER-BASED ACADEMIC ASSESSMENT SYSTEM (CAAS)

Researchers associated with the Laboratory for the Assessment of Academic Skills, at the University of Massachusetts, Amherst are in the process of developing a broad based assessment system called the Computer-based Academic Assessment System or CAAS. The CAAS system is designed to assess reading and math component processing skills in a manner that will prove to be both diagnostic and suggestive of remedial prescriptions.

The approach to the assessment of reading skills described below contains many of the elements suggested by the preceding review of cognitive developmental theory. However, the system is still under development and there are important features not included which are currently being developed and which will be added to the system in the future.

The CAAS system includes measures of listening and reading comprehension and measures of the speed and accuracy of performance on a number of tasks designed to measure cognitive components that support skilled reading.

Comprehension Tests

CAAS assessment begins with the administration of Sentence Verification Technique (SVT) listening and reading comprehension tests. Royer and his associates (1990; Royer, Carlo, and Cisero, 1992) have presented evidence from over 10 years of research that the SVT is a reliable and valid technique for assessing a reader's ability to understand a particular text.

The SVT listening and reading comprehension tests used in the study to be reported were based on materials from the same sources, thereby providing reasonably parallel measurement in both modalities. The tests are a critical component of the CAAS system for two reasons. First, they provide a means of assessing comprehension skills alone, apart from more general intellectual skills that are typically assessed with standardized reading tests (Johnston, 1984). Second they provide a means for determining if a comprehension problem is a specific reading problem.

Several authors (e.g., Hall and Humphreys, 1982; Stanovich, 1986a, b) have written of the importance of separating specific reading problems from more general cognitive problems. Stanovich (1986a) places greater emphasis on this point. He writes:

> . . . one assumption that is essential to all definitions of reading disability is the assumption of specificity. This assumption underlies all discussion of the concept, even if it is not stated explicitly. Simply put, it is the idea that a child with this type of learning disability has a brain/cognitive deficit that is reasonably specific to the reading task. That is, the concept of a specific reading disability requires that the deficits displayed by such children not extend too far into other domains of cognitive functioning. If they did, there would already exist research and educational

> designations for such children (low intelligence, slow learner, etc.), and the concept of reading disability would be superfluous. (p. 384)

The administration of both listening and reading SVT tests respond to the specificity issue by sorting students who have poor reading comprehension performance into two categories. The first category contains readers who perform adequately on the listening tests but poorly on the reading tests. These are children who are likely to have a specific reading disability because more general cognitive problems would result in poor performance on both the listening and the reading tests. The second category, occupied by children who exhibit poor performance on both the listening and reading tests, could consist of two types of children: Those who have poor general cognitive skills, and those who do not have the prior knowledge necessary to understand the content of both the listening and reading tests. It is beyond the scope of this paper to discuss ways of differentiating between these two types of students, but it should be apparent that listening comprehension assessments based on material varying in content should differentiate between the two problem categories.

The CAAS system does not use reading and listening SVT performance for any purpose other than identifying students with specific reading disabilities. However, it should be noted that patterns of performance on different item types comprising the SVT tests may also have diagnostic potential. Carlisle (1989a, b, 1990) has reported a number of studies in which students with well defined reading problems can be sorted into meaningful categories based on differing patterns of item performance on SVT listening and reading tests.

Computer Administered Tasks

The next step in CAAS assessment is the administration of a battery of computer-based tasks that are designed to measure component processes of reading. The following tasks are included in the computer battery: a simple response time task, the Posner letter match task (e.g., Posner, Boies, Eichelman, and Taylor, 1969), naming of real and pseudowords (measured by vocalization latency), a category match task (deciding if two words belong to the same category), and two variations of a cloze task designed to measure syntactic processing and sentence comprehension. Examples of items in each of the tasks in the computer battery are presented in Table I and a more complete description of each of the tasks follows.

Table I. Tasks and Sample Stimuli for CAAS Computer Battery

Task	Sample stimuli
Posner letter-match task	AA, bb, DD, ee
1. Physical identity	Aa, Bb, Dd, Ee
2. Name identity	AB, ba, DE, Gh
3. Different	
Word naming	
1. Three-letter words	you, are, but, one
2. Four-letter words	show, goes, jump, help
4. Five-letter words	table, story, horse, cover
5. Six-letter words	dinner, banner, carton, fought
Pseudoword naming	
1. Three-letters	baw, ret, teg, bex
2. Four-letters	pold, nast, soat, stip
4. Five-letters	nable, frint, glope, flide
5. Six-letters	neeper, sorbid, nourly, larton
Concept activation	
1. Matches	car/truck, arm/leg, socks/shirt
2. Mismatches	bus/stool, nose/apple, bird/bed
Syntactic analysis	These two flowers are/is mine.
	Mary was/were writing a letter.
	She ate three slices of/at cheese.
Semantic analysis	The boy drank/ate his milk.
	The school bell rang/blew.
	The farmer planted/played the corn.

Simple Response-Time Task. This task measures the speed and accuracy of responding to non-verbal stimuli (*** or +++) appearing on the computer screen. The task provides an index of physical dexterity and response speed that could be used as a covariate or as a point of consideration in performance analysis and interpretation. In addition, this task is presented first in the computer battery and serves to acclimate the examinee to the task demands of the system.

Letter-Level Task. The Posner letter match task involves having subjects decide if two letters have the same name (e.g., Aa—same, Ab—different). This task was included as a measure of letter-identification skill (physical identity condition—aa or AA) and speed of access to long-term memory (name identity condition—Aa). Letter identification speed has been shown to increase with age and ability (Doehring, 1976; Biemiller, 1977–1978). Speed of access to long-term memory has proven to be a reliable discriminator between skilled and less-skilled beginning readers (Chabot *et al.*, 1983) as well as adult readers (Hunt, Lunneborg, and Lewis, 1975; Jackson and McClelland, 1979; Palmer, Macleod, Hunt, and Davidson, 1985).

The latest version of the CAAS system provides the option of substituting a letter naming task for the Posner task. The letter naming task involves the presentation of an upper or lower case letter on the computer screen. The examinee says the name of the letter into a microphone.

Word-Level Tasks. Vocalization latencies for naming of words and pseudowords were included as indices of decoding ability. Word-identification skills are assessed with the vocalization latencies to real words. The ability to identify words is considered an important individual difference parameter (LaBerge and Samuels, 1974; Curtis, 1980; Stanovich, 1980; Perfetti, 1985). However, even after words can be recognized with perfect accuracy, word-recognition times continue to decrease with age (Biemiller, 1977–1978). The pseudowords were included because the task of identifying pseudowords is considered a test of phonological recoding or the ability to apply grapheme-phoneme correspondence rules (De Soto and De Soto, 1983). Performance on naming tasks involving pseudowords has been shown to reliably discriminate between skilled and less-skilled readers (Perfetti and Hogaboam, 1975; Hogaboam and Perfetti, 1978).

Concept Activation. Speed of access to semantic memory is assessed in the CAAS system using a category match task. Speed and accuracy of semantic memory processes have been shown to vary as a function of reading ability in beginning readers (Perfetti and Lesgold, 1977; Chabot *et al.*, 1983). Chabot *et al.* argue that a category match task (deciding if two words belong to the same category) is a measure of both semantic memory access and search processes.

Sentence Processing. Finally, application of syntactic and semantic knowledge was assessed with two tasks that are variations of the cloze procedure. In the syntax task, subjects were asked to select a word to complete a blank in a sentence. The choices varied in syntactic correctness (i.e., tense, verb agreement). Using a similar procedure, Guthrie (1973) showed that this task discriminated between skilled and less-skilled readers. In the sentence comprehension task, subjects were also presented with a sentence containing a blank and were asked to choose a word to fill in the blank. In this task, the choices varied in semantic correctness. This variant of the cloze procedure (Brickley, Ellington, and Brickley, 1970) was considered to be a measure of sentence comprehension.

Although these are only a sampling of the possible tasks that might be used to examine component reading skills, they do represent several processing skills at several levels that have been demonstrated to distinguish between readers of differing abilities.

RESEARCH ASSESSING THE DIAGNOSTIC POTENTIAL OF THE CAAS SYSTEM: AN EXAMINATION OF THE RELIABILITY, VALIDITY, AND THEORETICAL CONSISTENCY OF THE CAAS SYSTEM

The first study examining the diagnostic potential of the CAAS system entailed administering SVT tests and the computer tasks in May of 1989 to 112 students enrolled in grades 2 through 5. One year later the CAAS battery was readministered to 59 of the students. The data from these administrations were examined to determine if: (1) CAAS performance provided reliable and valid indices of reading performance, and if (2) CAAS performance provided results consistent with cognitive developmental theory.

Description of the Study. The subjects were 39 grade 2 students, 27 grade 3 students, 37 grade 4 students, and 9 grade 5 students (112 students total) who were enrolled in two elementary schools in a small Western Massachusetts school system. One year after the initial assessment, a follow-up study was conducted. We requested permission to retest all students who had participated in the initial phase of the study. Nineteen grade 3 students (note these were grade 2 students a year earlier), 13 grade 4 students, 19 grade 5 students, and 8 grade 6 students (59 total) returned their permission forms and were retested on the computer tasks.

All computer-administered tasks were presented on a Toshiba T3100/20 laptop computer. The stimuli were presented on the computer monitor. Subjects responded either by pressing buttons on a keyboard overlay, or by naming a word or pseudoword which was picked up by a microphone that in turn activated a voice key. All of the computer tasks were programmed using Micro Experimental Laboratory software (Schneider, 1988). It should be noted that the CAAS system has now been programmed in its own software environment, a change that makes the system much easier to use for the specialized purpose of reading diagnostics.

Initial testing was conducted in two sessions. In the first session second grade subjects were administered only listening SVT tests because appropriate reading materials for the development of reading SVT tests were not available. Students in grades 3, 4, and 5 were administered both listening and reading tests during the first session. The tests were administered in group settings in which students initially listened to tape recorded instructions relevant to the listening tests, paused for questions, and then listened to audiotaped versions of the listening SVT tests. For students in grades 3, 4, and 5, the listening tests were immediately followed by reading SVT tests that were presented in booklet form. In the second session, subjects were individually administered the computer tasks.

Subjects were tested individually on the computer. After a brief set of introductory comments to inform them of the purpose and nature of the tasks, subjects completed the simple response time task, followed by presentation of the letter task, the word naming tasks, the concept activation task, and the sentence-processing tasks. If, during the computer session, it became apparent that the task demands exceeded the capability of the student, testing was terminated. Instructions for each task appeared on the screen and were read aloud by the experimenter prior to beginning the task. Each task was preceded by a set of practice trials. After the practice trials the experimenter answered any questions and reviewed the instructions as necessary for each subject. In all tasks, the stimulus remained on the screen until the subject made a response. Subjects received feedback regarding the accuracy of their responses on every trial. Both the speed and accuracy of every response were recorded by the computer (see Sinatra and Royer, 1993 for a more complete description of the computer testing procedures).

The Follow-up Study was conducted one year after the initial assessment. The procedures were nearly identical to those used in the first assessment. The exception was that the listening and reading SVT tests were not administered a second time.

At the same time that students were being assessed during both Study 1 and the Follow-up Study, teachers were asked to fill out questionnaires indicating the grade level of the basal series in which students were currently reading. In addition, during the time the Follow-up assessments were being collected, teachers were asked to rate children as being either Excellent, Good, Average, Below Average, or Poor readers for their age group. In some instances, teachers did not provide either the reading book level or the ratings of reading performances.

Results

Reliability and Validity of the CAAS System. If the CAAS system has utility as a diagnostic instrument, it must first be shown that performance on the CAAS battery is both a reliable and valid measure of reading skill. The reliability and validity of the SVT technique has been established previously for both reading and listening comprehension (Royer, Kulhavy, Lee, and Peterson, 1986; Royer, 1990; Royer *et al.*, 1992). For the present study, the reliability of the SVT tests was determined using Cronbach's Alpha and averaged .75 for the listening tests and .80 for the reading tests.

Indices of the reliability of the response-time measures contained in the CAAS battery were obtained by estimating the components of variance attributable to subjects and to stimulus items for each of the tasks. (For a discussion of obtaining an estimate of the consistency of measurement of items from the components in an analysis of variance see Myers, 1979). The reliability indices derived from this procedure ranged from .88 to .97 with an average of .94 and support the conclusion that the response time measures were reliable.

The next set of analyses were concerned with evaluating one aspect of the validity of the computer-administered tests. A necessary condition for validating component processing performance as an index of reading skill is establishing that performance on the component tasks contained in the computer battery is, in fact, related to reading performance. To accomplish this purpose, performance on the computer tasks was related to three indices of reading skill: grade level, students' current reading book level, and teacher ratings of reading skill.

The general form of the analyses that related indices of reading skill to computer task performance was to use the skill indices as independent variables, task performance as dependent variables, and Bonferoni *t*s with family-wise error rates controlled at 5% as the statistical criteria for significance. The first set of analyses used grade as an independent variable and accuracy and response time as dependent variables. These analyses indicated that response accuracy, which averaged over 90% correct for each of the tasks, generally did not vary as a function of grade. The exception was the word and pseudoword naming tasks where accuracy significantly improved as a function of grade. In contrast to the accuracy analyses, those analyses using response time as dependent variables indicated that response time varied significantly as a function of grade. The nature of the variation was that response time decreased as grade increased.

A second procedure for examining the relationship between computer task performance and reading skill involved using categories of reading book levels as the independent variable. Basal reading book levels were obtained for 95 of the students in the study. Reading book levels indicated the grade and book level at which a particular child was currently reading (for example, 2.1 indicates a child is reading at Grade 2, book 1). Within each grade, these levels were collapsed into three groups (low, medium, and high) to provide an index of reading ability. This resulted in groups consisting of 11 subjects in each of the three groups in Grade 2; 6 in each group in Grade 3; and 9, 11, and 6 in each of the three groups in Grade

4. Grade 5 did not participate in this analysis because of the small number of students in Grade 5.

Separate ANOVAs were calculated using reading book level as a grouping variable and response time on each of the computer tasks as the dependent measure. The Posner letter match task did not vary significantly as a function of ability groups. This result is consistent with the expectation that by second grade most subjects will have mastered letter identification.

Whereas the Posner task did not vary significantly across ability groups, the word and pseudoword naming times did tend to vary systematically with reading ability for Grades 2 and 3. Of the higher-level tasks, only sentence comprehension significantly discriminated between ability groups in Grades 2, 3, and 4. For Grade 3, in addition to the sentence-comprehension task, the category match and syntactical analysis task also discriminated between ability groups. It should also be noted that even the tasks that did not reach levels of statistical significance with a strict family-wise error rate varied in a manner such that students reading in lower level reading books tended to respond slower on the component processing tasks than did students reading in higher-level reading books.

As was the case with previous analyses, analyses conducted using accuracy scores as a dependent variable indicated there was little variability associated with variations in reading book level and task. The reason for this is that most tasks had average accuracy scores above .9. The only tasks to show any significant differences between ability groups in accuracy of performance were the naming tasks.

We sought to obtain further evidence for the validity of the CAAS system by sorting examinees into categories based on their CAAS system performance and then examining independent indices of reading competence as a function of these categories. The separation of the students into CAAS performance categories was accomplished using the SPSS/PC+ cluster analysis procedure (Norusis, 1988). For those unfamiliar with cluster analysis, it entails creating standardized scores and then computing the squared Euclidean distance which is the sum of the square differences over all of the variables. Cases can then be grouped by clustering together cases having small Euclidean distances. The process we followed entailed entering Posner, Word Naming, Pseudoword Naming, Category, Syntax, and Semantic data into the statistical routine and selecting the solution involving three clusters. The cluster analyses were performed separately for each grade. After the cluster analysis procedure grouped the students, High, Medium, and Low labels were attached to each of the three clusters based on a visual inspection of the cluster scores.

Table II shows the average performance of the cluster groups on the SVT listening and reading tests, and it indicates the grade level of the reading book that the students were reading in at the time of Study 1 and at the time of the Follow-up Study. The table shows that the cluster groups were nearly perfectly ordered on the Listening SVT tests, the Study 1 reading book levels, and the Follow-up Study reading book levels. The data for the reading SVT tests were not as nicely ordered in that the Medium cluster group tended to perform slightly better on the tests than the High cluster groups. However, both the High and the Medium cluster groups outper-

Table II. Mean Listening and Reading Comprehension Indices for Average, High, Medium, and Low Cluster Readers in Grades 2–5

Measure and cluster	Grade 2	Grade 3	Grade 4	Grade 5
SVT listening comprehension (percent correct)				
High	86.6	84.3	91.4	93.7
	(26)[a]	(3)	(2)	(1)
Medium	86.2	84.5	84.2	85.9
	(5)	(20)	(32)	(7)
Low	71.8	80.1	75.0	82.8
	(4)	(3)	(2)	(1)
SVT reading comprehension (percent correct)				
High	—[b]	77.0	79.6	85.9
	—	(3)	(2)	(1)
Medium	—	80.9	84.6	87.0
	—	(20)	(32)	(7)
Low	—	70.3	67.9	29.6
		(3)	(2)	(1)
Reading book level: Study 1 (in grade level units)				
High	2.7	5.2	6.2	8.1
	(26)	(2)	(1)	(1)
Medium	2.0	4.2	5.2	6.5
	(5)	(17)	(24)	(7)
Low	1.3	2.8	4.2	3.3
	(2)	(3)	(1)	(1)
Reading book level: Follow-up Study (in grade level units)				
High	4.5	5.9	6.0	—[c]
	(2)	(2)	(1)	—
Medium	4.0	5.1	6.0	—
	(3)	(9)	(22)	—
Low	2.8	4.0	5.0	—
	(2)	(2)	(1)	—

[a]Number of subjects contributing to mean.
[b]Second grade students did not take the reading SVT tests.
[c]Fifth grade data was not obtainable.

formed the Low cluster groups on the reading SVT tests by a considerable margin.

Table III shows the final item of data relevant to comparing the cluster groups. At the time of the Follow-up Study teachers were asked to rate the current reading ability of participating students. As can be seen in the table, teacher rating of reading competence tended to vary in accordance with clusters formed on the basis of data collected one year before the ratings were made. In short, CAAS assessments obtained a year earlier were good predictors of teacher ratings of reading competence collected the following year.

Consistency of CAAS Performance with Cognitive Developmental Theory. Sinatra and Royer (1993) report several analyses indicating that the patterns of performance on the CAAS system were consistent with cognitive developmental theory. To briefly summarize their results, they found that the word identification skills of older students (grade 3 and 4) are more strongly related to sentence comprehension than the word identification skills of younger students (grade 2). This is consistent with the cognitive developmental model reviewed earlier (particularly Perfetti's view, i.e., 1992). Specifically, young readers' word-identification processes are relatively slow and not yet encapsulated and automatic. Readers at this stage rely on contextual information to identify words. Thus, cognitive developmental theory would predict that there would be a relatively small

Table III. Number of Students Receiving Teacher Ratings Obtained 1 Year After CAAS Assessments as a Function of High, Medium, and Low CAAS Clusters and Grade

	Teacher rating				
Cluster	Excellent	Good	Average	Below average	Poor
Grade 2					
High	16	6	—	—	—
Medium	—	4	—	—	—
Low	—	—	—	1	1
Grade 3					
High	3	—	—	—	—
Medium	6	5	3	—	—
Low	—	—	2	1	—
Grade 4					
High	2	—	—	—	—
Medium	9	10	5	1	—
Low	—	—	—	1	—
Grade 5					
High	1	—	—	—	—
Medium	3	2	—	—	—
Low	—	—	—	1	—

relationship between young readers' (e.g., second graders) word-identification skills and their sentence comprehension when compared to the same relationship for third and fourth grade students. However, more mature readers have developed proficient word identification skills through reading and experience with text. Their word identification skills are becoming modularized and automatic. At this stage, word identification makes a stable and consistent contribution to sentence comprehension. The third and fourth graders in the Sinatra and Royer (1993) study had presumably reached that stage. On the whole, the results presented by Sinatra and Royer (1993) are consistent with the cognitive developmental theory outlined in the introduction to the paper.

The Identification of Readers with Specific Reading Disabilities

The third criterion that we suggested for a reading diagnostic system was that it should supply specific information about the nature of the reading problem that a reader was experiencing. A study seeking support for this criterion (Cisero, Royer, Merchant, and Wint, 1994) involved administering the CAAS battery to students having known reading disabilities and then examining the data for evidence of different patterns of performance as a function of the disability.

The subjects in the study were 68 students enrolled in a college in Western Massachusetts. Forty of the students were nondisabled readers recruited from an introductory psychology class. The remaining 28 were recruited through the Learning Disabilities Program at the college. Twelve of the students were identified as dyslexic and the remaining 16 were identified as having an undifferentiated learning disability. All of the students were administered the computer-based tasks in the CAAS system. The tasks were the same as those used with elementary school children.

The data obtained from the study were subjected to an multivariate analysis of variance performed separately on accuracy and response time scores. These analyses utilized group (nondisabled, dyslexic, LD) as a between subject factor and type of task as a within subject factor. The analysis of the accuracy scores showed a main effect for group with the nondisabled group having the highest overall accuracy score (96%), followed by the LD students (95%) and the dyslexic students (93%). It should be noted that in this analysis, and in the one to follow, there was also a significant effect for task, though this result is not surprising (particularly in the analysis of response time).

The most interesting results, however, were obtained in the analysis of response time where there was both a significant group main effect and a significant group by task interaction. This result is depicted in graphic form in Fig. 1. The response time data contributing to Fig. 1 was transformed in several ways to allow an easy visual analysis of the patterns of data for the three groups. These transformations involved calculating means and standard deviations for each of the tasks by group. The task performance of the learning disabled (LD) and dyslexic groups were then converted into effect sizes by subtracting the task mean from the task mean of the nondisabled group and then dividing by the standard deviation of the nondisabled group. This conversion, in effect, provides a Z score indication of where the average LD or dyslexic student would be placed in the distribution *if* they were in the nondisabled group. The following step in the transformation process was to convert the effect size scores into percentile scores based on area under the normal curve encompassed by the effect size Z scores. In this process, nonlearning disabled students were defined as being at the 50th percentile (a defining feature of creating effect size scores), and the performance of the other groups was depicted in relation to their performance.

Figure 1 presents the results of these transformation processes. The reader will note that there are striking differences between the patterns of performance for the LD and the dyslexic groups. The LD group is consis-

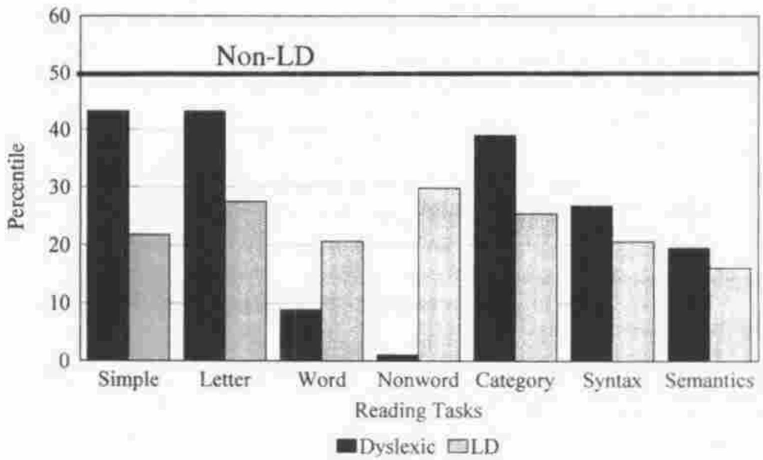

Fig. 1. Performance of dyslexic and learning disabled college students on reading tasks as compared to non-learning disabled college students. Non-LD students are represented by the dark line at the 50th percentile.

tently slower than the non-LD group on all of the tasks (including the simple response time task which does not involve reading), performing somewhere around the 20th percentile on each task. This pattern of performance is suggestive of a general cognitive "sluggishness."

In contrast, the dyslexic group generally performed better than the LD group on all tasks (though somewhat below the non-LD group), with two notable exceptions: the word and pseudoword naming tasks. Performance on both of these tasks is below the tenth percentile relative to the performance of the non-LD readers, and in the case of pseudoword naming, the response times are even well below those for the LD group. This pattern of performance suggests a specific deficit associated with the identification of words, and particularly severe problems in utilizing phonemic information to assist in this process.

DISCUSSION

In the introduction to this article we suggested that cognitive developmental theory could provide the foundation for a theory-based approach to reading diagnostics. Furthermore, we suggested that four conditions would have to be met before the viability of a diagnostic approach based on cognitive developmental theory could be established. First, the diagnostic technique would have to provide reliable assessments, and it would have to have been shown to be a valid means of assessing reading performance. Second, assessments provided by the technique would have to be consistent with expectations derived from cognitive developmental theory. Third, assessments provided by the technique would have to identify specific deficiencies that could be the source of the reading difficulties that the reader is experiencing. Fourth, assessments provided by the technique would have to lead to prescriptive procedures that alleviate the reading difficulties being experienced. We will discuss the CAAS system as it relates to each of these criteria.

Condition 1: Reliability and Validity of the CAAS System

Previous research (Royer *et al.*, 1986, 1992) has determined that SVT listening and reading comprehension tests provide reliable assessment. The reliabilities for the SVT tests obtained in this study were somewhat lower than those obtained in previous research, though they were still in the acceptable range. A possible reason for the lower reliabilities was that there was an unusual number of zero variance items on several of the test forms (all subjects got the item correct), thereby effectively reducing the length of the tests with a corresponding lowering of test reliability.

The analyses indicated that the reliabilities of all of the computer-administered tasks used in the study were excellent. Moreover, the data suggested that performance on the computer tasks were valid indices of reading competence. This was demonstrated in three separate ways. First, the analyses showed that response-time performance on the computer tasks in most cases reliably changed as a function of grade. The nature of this change was that task performance improved as grade increased. A second set of analyses examined computer-task performance within grade as a function of level of the reading books assigned to students. These analyses showed that students who were reading in advanced reading books were generally faster on the computer tasks than were students who were reading in lower level reading books.

The third set of analyses involved dividing students into High, Medium, and Low clusters based on their CAAS performance and then examining the performance of students in these clusters on other indices of reading competence. Examinations of mean performance indicated that the CAAS based groupings differed in listening and reading comprehension. They also differed in the level of the reading book the students were reading in when the Study 1 data were collected and when the Follow-up Study was conducted. The data also indicated that students in the cluster groups differed in teacher evaluations of reading competence collected during the Follow-up Study.

It should be noted that these relationships were obtained when response time was the dependent variable. In contrast, those analyses using response accuracy as dependent variables showed little variation as a function of indices of reading skill. The likely reason for failing to find effects on response accuracy is that most of the readers in the study were relatively skilled and, thus, highly accurate. In addition, the level of difficulty of the stimulus was targeted at approximately the third grade level. If the study had included a reasonable number of poor readers it is likely that response accuracy would have been sensitive to reading skill. But the fact that response speed differentiated between reading skill whereas response accuracy did not testifies to the power of response speed as an index of reading skill. Taken as a whole, these results satisfy the first criteria for a diagnostic system. That is, there is evidence supporting the reliability and validity of the CAAS system.

Condition 2: Consistency of the CAAS Findings with Cognitive Developmental Theory

The Sinatra and Royer (1993) study reported data indicating that the results of CAAS assessments varied in accordance with expectations derived from cognitive developmental theory of reading. Their results, along with data reported here, will have to be verified in further research, but at this point they are certainly consistent with the principles of cognitive developmental theory.

Condition 3: Identification of Specific Reading Deficiencies

The Cisero *et al.* (1994) study shows clearly differentiated patterns of performance for nondisabled college readers, readers with an undifferentiated learning disability, and readers who are dyslexic. These differential patterns of performance support the ability of the CAAS system to differentiate between readers having various types of disabilities, though further research involving larger numbers of subjects will have to verify this support.

Condition 4: Prescriptive Potential of the CAAS Assessment

To this point we have not conducted research evaluating the ability of the CAAS system to provide useful prescriptive information. There are, however, clear signs that the CAAS system might provide such information. One common denominator among all of the readers having below average performance in our studies was that their word-identification skills were poorer than the skills of their more successful cohorts. An obvious implication of this finding is that instructional procedures designed to improve word identification skills could potentially breach a barrier that had blocked the development of reading ability.

Recent theoretical and empirical work in word identification has led to the development of an instructional strategy called "decoding by analogy" (Gaskins *et al.*, 1991). Teachers can use this strategy to help children identify unfamiliar words by looking for spelling patterns or "rimes" (such as recognizing *ail* and *ate* in the word tailgate). Research indicates that decoding by analogy may be more efficient than traditional word identification strategies because spelling patterns are more consistent across words than individual letter-sound correspondences (Adams, 1990; Cunningham, 1975–1976). In fact, recent research has shown the effectiveness of this instructional approach for the development of children's word recognition skills (Dewitt, Snyder, and Coressel, 1992; Wagstaff and Sinatra, 1994).

The results of the current study imply that an instructional approach designed to improve word identification ability should focus on both the speed and accuracy of performance. Many of our Low-cluster, elementary-school readers were inaccurate in the naming tasks. There were a few, however, that were near perfect in accurately naming words or pseudowords, but they did so much more slowly than their more successful classmates. Computer gaming techniques (e.g., Frederiksen *et al.*, 1985a, b; Olson and Wise, 1992) are examples of instruction that could be directed at improving both the speed and the accuracy of word-identification skills.

The fact that all of the Low cluster readers in our study had poor word identification skills, does not, of course, mean that poor word identification skills are the source of all reading difficulties. It could be that readers have difficulty because of inadequate letter identification skills or poor skills associated with concept activation or syntactic and semantic analysis.

The studies reported in this article did not entail any effort to repair probable sources of reading difficulty identified by the CAAS system, therefore no claim of meeting the fourth criteria for a reading diagnostic system can be made. However, as the discussion in the preceding paragraph indicates, there are several avenues for pursuing support for this criteria that can be the subject of future research.

Limitations of the System

The data we report in this paper are limited in that they focus on reading skills that range in complexity from the identification of letters to listening and reading comprehension. Moreover, it should be noted that the comprehension assessed by the SVT procedure entails an assessment of whether a reader can extract the surface meaning of a text. SVT tests, by design, do not measure an examinee's ability to draw inferences from text or to think about material that has been read.

The CAAS system also does not provide assessments of phonemic awareness, metacognitive or strategic abilities, or prior knowledge. However, we are in the process of adding a phonological awareness task to the CAAS system and are evaluating a spelling pattern recognition task that may be important for determining level of word recognition development (Ehri, 1991).

ACKNOWLEDGMENTS

We would like to thank Donald Compton, Joanne Carlisle, Barbara Greene, Jeff Walczyk, and Steve Benton for comments that have greatly improved the quality of our article.

REFERENCES

Aaron, P. G. (1991). Can reading disabilities be diagnosed without using intelligence tests? *J. Learning Disab.* 24(3): 178-186.

Adams, M. J. (1990). *Beginning to Read: Thinking and Learning About Print*, MIT Press, Cambridge, MA.

Anderson, J. R. (1983). *The Architecture of Cognition*, Harvard University Press, Cambridge, MA.

Anderson, R. C., and Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension, In Pearson, P. D., Barr, R., and Kamil, M. (eds.), *Handbook Reading Research*, Longman, New York.

Anderson, R. C., Reynolds, R. E., Shallert, D. L., and Goetz, E. T. (1977). Frameworks for comprehending discourse. *Am. Educ. Res. J.* 14: 367-381.

Allington, R. L. (1984). Content coverage and contextual reading in reading groups. *J. Reading Behav.* 16: 85-96.

Baker, L., and Brown, A. L. (1984). Metacognitive skills in reading. In Pearson, P. D. (ed.), *Handbook of Reading Research*, Longman, New York, pp. 353-395.

Ball, E. W. and Blachman, B. A. (1988). Phoneme segmentation training: Effect on reading readiness. *Ann. Dyslexia* 38: 208-225.

Bejar, I. I. (1984). Educational diagnostic assessment. *J. Educ. Measurement* 21(2): 175-189.

Biemiller, A. (1977–1978). Relationship between oral reading rates for letters, words, and simple text in the development of reading achievement. *Reading Res. Quart.* 13: 223-253.

Bradley, L. (1988). Rhyme recognition and reading and spelling in young children. In Ellis, W. (ed.), *Intimacy with Language: A Forgotten Basic in Teacher Education* (2nd ed.), Orton Dyslexia Society, Baltimore.

Bradley, L., and Bryant, P. E. (1983). Categorizing sounds and learning to read—A causal connection. *Nature* 30: 419-421.

Bradley, L., and Bryant, P. E. (1985). *Rhyme and Reason in Reading and Spelling*, University of Michigan Pres, Ann Arbor.

Brown, A. L., and Campione, J.C. (1986). Psychological theory and the study of learning disabilities. *Am. Psychol.* 41: 1059-1068.

Bryant, P. E., and Bradley, L. (1985). *Children's Reading Difficulties*, Blackwell, Oxford.

Brickley, A. C., Ellington, B. J., and Brickley, R. T. (1970). The cloze procedure: A conspectus. *J. Reading Behav.* 2: 232-239.

Carlisle, J. F. (1989a). The use of the Sentence Verification Technique in diagnostic assessment of listening and reading comprehension. *Learning Disab. Res.* 5: 33-44.

Carlisle, J. F. (1989b). Diagnosing comprehension deficits through listening and reading. *Ann. Dyslexia* 39: 159-176.

Carlisle, J. F. (1990). Diagnostic assessment of listening and reading comprehension. In Swanson, H. L., and Keogh, B. (eds.), *Learning Disabilities: Theoretical and Research Issues*, Erlbaum, Hillsdale, NJ.

Chabot, R. J., Petros, T. V., and McCord, G. (1983). Developmental and reading ability differences in accessing information form semantic memory. *J. Exp. Child Psychol.* 35: 128-142.

Cisero, C. A., Royer, J. M., Marchant, H., and Wint, F. (1994). *Diagnostic Identification of Dyslexic and Learning Disabled College Students Using the Computer-Based Academic Assessment System (CAAS).* Paper presented at the Annual Meeting of the American Educational Association, New Orleans, LA.

Cunnigham, P. M. (1975–1976). Investigating a synthesized theory of mediated word identification. *Reading Res. Quart.* 11: 127-143.

Curtis, M. E. (1980). Development of components of reading skill. *J. Educ. Psychol.* 72: 656-669.

DeSoto, J. L., and DeSoto, C. L. (1983). Relationship of reading achievement to verbal processing abilities. *J. Educ. Psychol.* 75: 116-127.

Dewitt, P. A., Snyder, K., and Coressel, L. L. (1992). *Comparing an Analogy and Phonics Approach to Word Recognition*. Paper presented at the Annual Meeting of the National Reading Conference, San Antonio.

Doehring, D. G. (1976). Acquisition of rapid reading responses. *Monogr. Soc. Res. Child Dev.* 41(2, of Serial No. 165).

Durrell, D. D., and Catterson, J. H. (1980). *Durrell Analysis of Reading of Reading Difficulties*, Psychological Corporation, New York.

Ehri, L. C. (1991). Development of the ability to read words. In Barr, R., Kamil, M. L., Mosenthal, P. B., and Pearson, P. D. (eds.), *Handbook of Reading Research* (Vol. 2), Longman, New York, pp. 383-417.

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In Gough, P. B., Ehri, L., and Treiman, R. (eds.), *Reading Acquisition*, Erlbaum, Hillsdale, NJ, pp. 107-143.

Fodor, J. A. (1983). *The Modularity of the Mind*, MIT Press, Cambridge, MA.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In Cooper, W. E., and E. Walker, (eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, Erlbaum, Hillsdale NJ.

Frederiksen, J. R., Warren, B. M., and Rosebery, A. S. (1985a). A componential approach to training reading skills: Part 1. Perceptual units training. *Cognit. Instr.* 2: 91-130.

Frederiksen, J. R, Warren, B. W., and Rosebery, A. S. (1985b). A componential approach to training reading skills: Part 2: Decoding and the use of context. *Cognit. Instr.* 2: 175-205.

Gaskins, R. W., Gaskins, J. C., and Gaskins, I. W. (1991). A decoding program for poor readers—and the rest of the class, too! *Language Arts* 68: 213-225.

Glaser, R. (1981). The future of testing: A research agenda for cognitive psychology and psychometrics. *Am. Psychol.* 36: 923-936.

Glass, G. (1986). Testing old, testing new: Schoolboy psychology and the allocation of intellectual resources. In Plake, B. S., and Witt, J. C. (eds.), *The Future of Testing*, Erlbaum, Hillsdale, NJ, pp. 9-27.

Goodman, Y. M., and Burke, C. L. (1972). *Reading Miscue Inventory*, Macmillan, New York.

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7(1): 6-10.

Gough, P. B., Juel, C., and Griffith, P. L. (1992). Reading, spelling, and the orthographic cipher. In Gough, P. B., Ehri, L., and Treiman, R. (eds.), *Reading Acquisition*, Erlbaum, Hillsdale, NJ, pp. 55-48.

Guthrie, J. T. (1973). Models of reading and reading disability. *J. Educ. Psychol.* 65: 9-18.

Hall, J., and Humphreys, M. (1982). Research on specific learning disabilities: Deficits and remediation. *Topics Learning Disab.* 2: 68-78.

Hogaboam, T. W., and Perfetti, C. A. (1978). Reading skill and the role of verbal experience in decoding. *J. Educ. Psychol.* 70: 717-729.

Hunt, E., Lunneborg, C., and Lewis, J. (1975). What does it mean to be high verbal? *Cognit. Psychol.* 7: 194-227.

Jackson, M. D., and McClelland, J. L. (1979). Processing determinants of reading speed. *J. Exp. Psychol. General* 108: 151-181.

Johnston, P. (1992). *The Language of Assessment and the Assessment of Language.* Paper presented at the Annual Meeting of the National Reading Conference, San Antonio, TX.

Johnston, P. H. (1984). Assessment in reading. In Pearson, P. D. (ed.), *Handbook of Reading Research*, Longman, New York, pp. 147-182.

LaBerge, D., and Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognit. Psychol.* 6: 293-332.

Linn, R. L. (1986). Educational testing and assessment: Research needs and policy issues. *Am. Psychol.* 41: 1153-1160.

Lundberg, I., Frost, J., and Peterson, O. P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Res. Quart.* 23: 263-284.

Mestre, J. P., and Royer, J. M. (1991). Cultural and linguistic influences on Latino testing. In Keller, G., Beneen, J., and Magallan, R. (eds.), *Assessment and Access: Hispanics in Higher Education,* State University of New York Press, Albany, NY, pp. 39-66.

Myers, J. L. (1979). *Fundamentals of Experimental Design* (3rd Ed.), Allyn and Bacon, Boston, MA.

Nagy, W. E., and Anderson, R. C. (1984). How many words are there in printed school English. *Reading Res. Quart.* 20: 233-253.

Norusis, M. J. (1988). *SPSS/PC+ Advanced Statistics V2.0*, SPSS Inc., Chicago, IL.

Olson, R. K., and Wise, B. (1992). Reading on the computer with orthographic and speech feedback: An overview of the Colarado Remediation Project. *Reading Writing* 4: 107-144.

Palincsar, A. S., and Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cogn. Instr.* 1: 117-175.

Palmer, J., Macleod, C. M., Hunt, E., and Davidson, J. E. (1985). Information processing correlates of reading. *J. Mem. Lang.* 24: 59-88.

Paris, S. G., Cross, D. R., and Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *J. Educ. Psychol.* 76: 1239-1252.

Perfetti, C. A. (1985). *Reading Ability*, Oxford University Press, New York.

Perfetti, C. A. (1988). Verbal efficiency in reading ability. In Daneman, M., MacKinnon, G. E., and Waller, T. G. (eds.), *Reading Research: Advances in Theory and Practice*, Academic Press, New York, pp. 109-143.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In Gough, P. B., Ehri, L., and Treiman, R. (eds.), *Reading Acquisition*, Erlbaum, Hillsdale, NJ, pp. 145-174.

Perfetti, C. A., and Hogaboam, T. W. (1975). Relationship between single word decoding and reading comprehension skill. *J. Educ. Psychol.* 67: 461-469.

Perfetti, C. A., and Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In Just, M., and Carpenter, P. (eds.), *Cognitive Processes in Comprehension*, Erlbaum, Hillsdale, NJ.

Posner, G. J. Strike, K. A. Hewson, P. W., and Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Sci. Educ.* 66: 211-227.

Posner, M. I., Boies, S. J., Eichelman, W. H., and Taylor, R. L. (1969). Retention of physical and name codes of single letters. *J. Exp. Psychol. Monogr.* 79: 1-16.

Pressley, M., Goodchild, F., Fleet, J., Zajchowski, R., and Evans, E. D. (1989). The challenges of classroom strategy instruction. *Elem. School J.* 89: 301-342.

Rayner, K., and Pollatsek, A. (1989). *The Psychology of Reading*, Prentice-Hall, Englewood Cliffs, NJ.

Roth, S. F., and Beck, I. L. (1987). Theoretical and instructional implications of the assessment of two microcomputer word recognition programs. *Reading Res. Quart.* 22: 197-218.

Royer, J. M. (1985). Reading from the perspective of a biological metaphor. *Contemp. Educ. Psychol.* 10: 150-200.

Royer, J. M. (1990). The Sentence Verification Technique: A new direction in the assessment of reading comprehension. In Legg, S. M., and Algina, J. (eds.), *Cognitive Assessment of Language and Math Outcomes*, Ablex, Norwood, NJ, pp. 144-191.

Royer, J. M., and Carlo, M. S. (1991). Assessing the language acquisition progress of limited English proficient students: Problems and a new alternative. *Appl. Measurement Educ.* 4: 85-113.

Royer, J. M., and Cunningham, D. J. (1981). On the theory and measurement of reading comprehension. *Contemp. Educ. Psychol.* 6: 187-216.

Royer, J. M., Kulhavy, R. W. Lee, J. B., and Peterson, S. E. (1986). The sentence verification technique as a measure of listening and reading comprehension. *Educ. Psychol. Res.* 6: 299-314.

Royer, J. M., Carlo, M. S., and Cisero, C. A. (1992). School-based uses for the Sentence Verification Technique of measuring listening and reading comprehension. *Psychol. Test Bull.* 5: 5-19.

Schneider, W. (1988). Micro experimental laboratory. An integrated system for IBM PC compatible. *Behav. Res. Meth. Instr. Comp.* 20: 206-217.

Schwartz, S. (1980). *Measuring Reading Competence: A Theoretical-Prescriptive Approach*, Plenum Press, New York.
Sinatra, G. M., and Royer, J. M. (1993). The development of cognitive component processing skills that support skilled reading. *J. Educ. Psychol.* 85: 509-519.
Spiro, R. J. (1980). Constructive processes in prose comprehension and recall. In Spiro, R. J., Bruce, B. C., and Brewer, W. F. (eds.), *Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education*, Erlbaum, Hillsdale, NJ, pp. 245-278.
Stanovich, K. E. (1986a). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Res. Quart.* 21: 360-407.
Stanovich, K. E. (1986b). Cognitive processes and the reading problems of learning disabled children: Evaluating the assumption of specificity. In Torgesen, J., and Wong, B. (eds.), *Psychological and Educational Perspectives on Learning Disabilities*, Academic Press, New York, pp. 87-131.
Stanovich, K. E. (1988). Explaining the difference between the dyslexic and the garden-variety poor reader. The phonological-core variable-difference model. *J. Learning Disab.* 21(10): 590-604.
Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity, *Dev. Rev.* 10: 72-100.
Stedman, L. C., and Kaestle, C. E. (1987). Literacy and reading performance in the United States from 1880 to the present. *Reading Res. Quart.* 22: 8-46.
Sternberg, R. J. (1981). Testing and cognitive psychology. *Am. Psychol.* 36: 1181-1189.
Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)Consideration of context effects. *J. Verbal Learning Verbal Behav.* 18: 645-659.
Vinsonhaler, J. S., Weinshank, A. B. Wagner, C. C., and Polin, R. M. (1983). Diagnosing children with educational problems: Characteristics of reading and learning disabilities specialists, and classroom teachers. *Reading Res. Quart.* 28: 134-164.
van Dijk, T. A., and Kintsch, W. (1983). *Strategies of Discourse Comprehension*, Academic Press, New York.
Wagstaff, J. M., and Sinatra, G. M. (1994). *Promoting Efficient and Independent Word Recognition: A New Strategy for Readers and Writers*, Manuscript submitted for publication.
Wade, S. E., Schraw, G., Buxton, W. M., and Hayes, M. T. (1993). Seduction of the strategic reader: Effects of interest on strategy and recall. *Reading Res. Quart.* 28(2): 92-114.
Walberg, H. J., and Tsai, S. (1983). Matthew effects in education. *Am. Educ. Res. J.* 20: 359-373.
Wolf, M. (1984). Naming, reading, and the dyslexias: A longitudinal review. *Ann. Dyslexia* 34: 78-115.

Copyright of Educational Psychology Review is the property of Kluwer Academic Publishing and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.