5. The Assessment of Cognitive Factors in Academic Abilities

Stephen L. Benton

_Kansas State University_, benton@ksu.edu

Kenneth A. Kiewra

_Utah State University_

Follow this and additional works at: [http://digitalcommons.unl.edu/buroscogpsych](http://digitalcommons.unl.edu/buroscogpsych)

Part of the [Cognitive Psychology Commons](http://digitalcommons.unl.edu/cogpsych), and the [Educational Assessment, Evaluation, and Research Commons](http://digitalcommons.unl.edu/assessment)
Nearly 30 years ago, Lee Cronbach (1957) distinguished between the two disciplines of correlational psychology, which investigated naturally occurring individual variance in behavior, and experimental psychology, which examined the effectiveness of certain treatments on behavior. Essentially, correlational psychology examined individual differences using factor analytic techniques; whereas experimental psychology attempted to eliminate individual differences using appropriate interventions. Cronbach believed that these two disciplines should join together to promote aptitude-treatment interaction (ATI) research that would identify effective treatments for certain types of individuals. With this combined approach, different treatments could be prescribed for skilled and less skilled individuals.

The ATI research methodology had limited success, however, because of inconsistency in findings and because of difficulty in replicating some of the treatments (Tobias, 1985). In addition, results rarely revealed disordinal interactions (which indicate that treatments differentially affect those on the lower and the higher ends of the performance continuum). One explanation for the lack of disordinal interactions was that methods for identifying skilled and less skilled students on a given academic behavior were not far advanced (Tobias, 1985). What was needed were precise methods for measuring specific skills required for successful academic achievement.

Recent developments in cognitive psychology have provided more precise methods that may help to advance both ATI research and the field of measurement. Sternberg (1977), for example, has investigated the underlying cognitive processes in intellectual behavior using componential analysis. Essentially, componential analysis investigates the underlying components involved in task per-
formance. By specifying these components and the various combination rules one might employ, differences can be observed among individuals in the number of components utilized, the combination rules employed, the order of the component operations, the mode of processing (e.g., serial vs. parallel), and the time required to execute a component. This approach is more precise than the previous factor analytic approaches of correlational psychologists, because the latter measured only the end products of behaviors and not the components of mental organization (Vernon, 1970, p. 100).

A more recent trend in cognitive psychology has been to examine differences in mental abilities between experts and novices in particular subject matter areas such as math and reading. Such cognitive curriculum analysis can not only specify cognitions that distinguish experts and novices, but can promote the construction of tests in particular subject-matter areas that can diagnose the cognitive difficulties underlying performance. By identifying the underlying components involved in academic performance, differences may then be observed that allow for more precise measurement and more effective treatments.

The cognitive approach to assessment suggests that there are several factors that contribute to successful or unsuccessful academic behaviors. These generally interactive factors are the learner’s declarative knowledge, procedural knowledge, control processes, cognitive strategies, and metacognitive processes. Cognitive psychology has advanced to the point where it can offer tools for measuring these factors that may help to clarify the specific interventions that must be made. In many instances, however, the tools are still being developed and applied to specific academic areas so that subject-matter remediation can be more precise.

The purpose of the current chapter, then, is to investigate how these cognitive factors may be measured within the academic domains of reading, writing, mathematics, and science. What immediately follows is a brief overview of the cognitive factors and a description of how they may be assessed, in general. (A more detailed account can be found in Meyer, 1981.) Following that overview, methods for assessing these factors within the various academic domains are discussed.

**DECLARATIVE KNOWLEDGE**

Declarative knowledge refers to knowledge of facts and information. Several researchers stress the importance of having appropriate declarative knowledge for demonstrating expertise in problem solving or in higher-order thinking. This view is well supported by Resnick (1984), who purports that thinking can only be taught in knowledge-rich areas, and by proponents of artificial intelligence who now share the view that intelligent thinking is knowledge based (Minsky & Papert, 1974).
Experts and novices working in a particular area differ in both the quantity and the quality of their domain-specific knowledge. Expert chess players, for example, have more knowledge about chess positions than do chess novices (Chase & Simon, 1973). Not only do novices have less knowledge, but their knowledge is often not as hierarchically organized as is the knowledge of experts (e.g., Chi, Glaser & Reese, 1982). Novices also appear to have fewer links or pathways among their memory nodes, thereby decreasing the activation of related knowledge. This limits both encoding and retrieval processes, which, of course, are necessary for effective problem solving, memorization, and comprehension.

Experts not only have sufficient declarative knowledge about the particular domain of inquiry, they also have knowledge about the structure of knowledge that helps them to learn and to understand. Someone trying to comprehend a story about baseball, for example, not only needs baseball knowledge, but also knowledge about the structure of stories. As we hear or read a story, we expect the events of that story to correspond to our story schema that may include an introduction, a characterization, a resolution, and so forth. When we comprehend, we selectively modify the story’s events to conform to our current schema. Therefore, stories presented in a manner inconsistent with our story schema are more difficult to comprehend (Thorndyke, 1977). Differences between good and poor readers (discussed later in greater detail) are, in fact, often due to the readers’ immature story schemata. This is why Resnick (1985) maintains that meaning is as much within the learner as it is upon the printed page. Resnick therefore advocates that particular story schemata be taught, because adequate schemata have transfer value that will increase the likelihood of comprehension across content areas.

If content knowledge and knowledge about the structure of knowledge contribute to expertise, then researchers and educators need methods for assessing such declarative knowledge. Cognitive psychologists have recently provided the tools for such measurement. The cognitive approach to assessing declarative knowledge involves analyzing verbal knowledge into composite units and indicating the structure governing those units. A structure model of a person’s knowledge is represented in the form of a network or a tree, both of which indicate symbolically the major elements of a person’s knowledge and the relationships among those elements—much like a sentence grammar indicates the parts and relations within a sentence. One derives a story schema, for example, by first breaking a story down into simple sentences. Each sentence is then placed within one of the designated components of a story schema. Story schemata are thought to include information about setting, theme, plot, and resolution. Each of these structures can, in turn, be analyzed into component parts. A setting contains information about characters, location, and time; a plot contains various episodes further comprised of subgoals, attempts to reach subgoals, and
outcomes. Thorndyke (1977) has suggested, for example, that the following parsing rules or categories capture most sentences contained within a narrative:

- **Rule 1:** Story = Setting + Theme = Plot + Resolution
- **Rule 2:** Setting = Character + Location + Time
- **Rule 3:** Theme + Events + Goal
- **Rule 4:** Plot = Episodes
- **Rule 4a:** Episode = Subgoal + Attempts + Outcome
- **Rule 5:** Resolution = Event or State

The structural model developed for a particular story can then be compared to the actual recall performance of individuals who have read the story. Because individuals generally use story schemata when comprehending stories (Thorndyke, 1977), such a comparison can identify specific gaps in an individual’s knowledge about the topic and, perhaps more importantly, about the structure of that knowledge.

It appears that cognitive structures are formed and used in various areas. As examples, Kintsch (1974) has identified a schema structure for scientific reports, and Spilich, Vesonder, Chiesi, and Voss (1979) have developed schema structures for understanding radio broadcasts of baseball games. The task, then, is for cognitive psychologists and psychometricians to develop these sorts of schematic structures in other areas. Doing so can permit instructors to teach the particular story schemata relevant to a particular area, and can help instructional designers design instructional materials consistent with the organizational structure of a content area.

**PROCEDURAL KNOWLEDGE**

Effective learners not only have adequate declarative knowledge, but also procedural knowledge that assists them in using declarative knowledge (Resnick, 1976; Woods, Resnick, & Groen, 1975). Skilled math students, for example, do not simply learn or memorize countless solutions to math problems; they are able to solve novel problems such as 638 divided by 19 because they have learned higher-order procedures or rules for doing so.

Cognitive psychologists, interested in the procedures incorporated by the expert and the procedural errors made by the novice in a given field, have developed two similar methods for representing and assessing procedural knowledge. One method is called a program, which is a step-by-step list of actions to be taken; the other is a flowchart, which is a set of boxes and arrows used to represent the processes and decisions one makes when solving a problem. A process model for solving a particular type of problem is derived by observing several individuals solve problems of that nature, and by interviewing them
about their procedures. A program or flowchart that corresponds to apparent procedures is then generated. The validity of the process model is then tested by comparing the processes and performance of other individuals solving similar problems to the processes and performance designated by the model. This sort of analysis can successfully pinpoint the specific procedural error(s) that a student is making.

An important example of the use of a process model for determining procedural errors was offered by Brown and Burton (1978), who developed a computer program called BUGGY that identifies process bugs or errors in the solving of subtraction problems. This program can determine the precise procedural error a student makes when subtracting. For example, a learner may always subtract the smaller number from the larger regardless of which number is on top; or the student may have difficulty borrowing across zero. The BUGGY program, then, does not only specify the correctness of a student’s response, but, more importantly, it identifies the particular procedural error(s) made by the student. With this type of information, teachers can become more effective in teaching specific procedures, rather than waiting and hoping that students discover them.

Cognitive tools like the BUGGY program certainly have implications for educational measurement as well. More programs specific to particular academic areas need to be constructed for developing tests that assess procedural errors and that determine the procedures used by resident experts. The stage has especially been set in the area of mathematics (Groen & Parkman, 1972; Resnick, 1976), which depends heavily on procedural knowledge. The importance of procedural knowledge in mathematics and other academic areas are discussed in later sections.

CONTROL PROCESSES

Recent research has indicated that individuals who differ on intelligence tests (e.g., Sternberg, 1977) and on ability tests (e.g., Hunt, 1978) also differ in their information-processing capabilities. Earl Hunt and his colleagues (Hunt, 1978; Hunt, Frost, & Lunneborg, 1973; Hunt, Lunneborg, & Lewis, 1975) have developed or modified several tasks that distinguish the particular information-processing components (sensory memory, short-term memory, working memory, and long-term memory) and/or control processes (attention, rehearsal, chunking, manipulating information in working memory, encoding, and searching long-term memory) associated with individual differences in verbal ability. In general, their work has indicated that differences between high- and low-verbal individuals can be more precisely interpreted as cognitive information-processing differences. In particular, they found that high- and low-verbal individuals differ on cognitive tasks involving search speed through long-term memory, the hold-
ing capacity of short-term memory, and the speed of manipulating information in working memory.

Developmental research on human information processing suggests that individual differences on cognitive tasks are due more to the effective use of control processes than to differences in the cognitive machinery or memory structures among individuals. Apparently, structure or hardware changes (e.g., number of holding slots in short-term memory) are negligible after early childhood (see Chi, 1978; Harris, 1978). Cognitive processes, like rehearsal and encoding, are, however, modifiable within certain limits. In fact, even learning disabled students (e.g., Torgesen, 1977), and retarded students (Campione & Brown, 1977), have shown significant improvement in memory performance following brief instruction in using rehearsal. (Research on the developmental aspects of control process training is reviewed by Chi, 1976; and by Hagen & Stanovich, 1977.)

Although substantial experimental memory research indicates that those who remember more are apt to use control processes more often or more effectively (see, for example, Bransford, 1979), it is, of course, impossible for researchers to directly assess these processes. Because these memory processes are not amenable to direct assessment, some researchers have corroborated the inferred processes through self-report techniques or through observation. Torgesen (1977), for example, observed the mouthing of words to infer rehearsal, and has observed picture rearrangements to infer organization processes in working memory.

Although cognitive psychologists have largely determined that differences in intelligence and verbal ability are due to cognitive processes, there remain questions about exactly what those processes are and how to more objectively measure them. As the pioneering work of Hunt and Sternberg continues to be applied to specific academic areas, perhaps these issues can be more successfully addressed. Furthermore, only by looking at control processes in specific areas can we be sure of their relative effectiveness for determining expertise when other cognitive factors, such as subject matter knowledge, are also considered.

**COGNITIVE STRATEGIES**

Another factor involved in solving general ability or specific academic problems is the cognitive strategies incorporated by the learner. Cognitive strategies are thoughts that influence how learners select, acquire, organize, or integrate new knowledge. These strategies represent a plan of attack for achieving a designated goal. In determining the types of strategies people use in solving problems, cognitive psychologists have presented people with problems and have asked them to think aloud as they solved them. From these self-reports, psychologists have identified the strategies that humans use—often programming them into a computer—and have, then, tested the programmed strategies against actual
human performance. Through this method, psychologists have identified the strategic behaviors that distinguish experts from novices with regard to solving general ability problems or problems associated with particular content areas.

Although there are several documented general strategies for problem solving (e.g., means-end analysis, working forward, reasoning by analogy, and brainstorming), cognitive psychologists have focused investigations on means-ends analysis. Ernst and Newell (1969) and Newell and Simon (1972), for example, constructed a computer program called General Problem Solver (GPS), using the self-report procedures described earlier, that uses means-ends analysis. GPS solves problems, as do many humans, by first determining a problem space consisting of one’s goal state, starting state, and all possible solution paths. (It should be apparent that appropriately structured declarative knowledge is critical for constructing the problem space.) Second, a goal-directed search is made by searching planfully through the problem space. This planful search is executed through the cognitive strategy of means-ends analysis, which entails generating and solving subgoals necessary for achieving the terminal goal. (A more complete description of GPS and means-ends analysis can be found in Ernst and Newell, 1969.)

The investigation of problem solving in particular content areas seems critical, because problem solving appears to depend substantially on the declarative knowledge one brings to the situation. Expert chess players, for example, actually behave less strategically, in some cases, than do novices. Expert players often do not use means-ends analysis to reduce the gap between their current state and the goal state, as novices are apt to do. Instead, experts respond almost automatically to the problem situation; namely, the current positioning of the chess pieces (e.g., Feltovich, 1981; Newell & Simon, 1972). Perhaps this is because of the expert’s superior knowledge and experience with a variety of possible chess positions. Thus, it is believed that the major differences between experts’ and novices’ problem-solving abilities in a particular area are due to the following knowledge-derived factors: (a) Experts represent the problem more effectively than do novices; (b) experts have more subject-matter knowledge that is usually organized hierarchically; and (c) experts, because of their more rich and coherently structured knowledge, are able to hold more information in memory and thereby entertain several hypotheses at a time. Novices, however, operate in a piecemeal fashion, reacting to the latest cue and forgetting to consider previous information. (See Gagné, 1985, pp. 136-161 for a description of problem-solving factors that distinguish novices and experts.)

Although the literature has reported modest success in teaching general problem-solving strategies like means-ends analysis (e.g., Covington, Crutchfield, & Davies, 1966), it seems that knowledge in an area is critical for applying effective strategies. Therefore, cognitive psychologists should not only continue to advance the technology of systems like GPS to more closely simulate human problem representation and solution search, but should especially focus these
efforts in particular areas where the amount and the structure of knowledge affects strategic behaviors. Subsequent sections describe the use of cognitive strategies in knowledge-rich domains.

**METACOGNITION**

Metacognition refers to what a person knows about his or her cognitive processes and the ability to control these processes by planning, choosing, and monitoring. A learner with good metacognition engaged in problem solving would likely be aware of his or her procedural and declarative knowledge relative to the problem, and would call upon appropriate control processes and cognitive strategies when necessary. Furthermore, he or she would periodically monitor the current state relative to the goal state. Thus, metacognition allows the learner to orchestrate or to control the cognitive factors previously discussed.

There is abundant research indicating developmental differences in metacognition between children of different ages. (See Brown, 1978, for a review.) Younger children, for example, are often unaware of their own knowledge relative to older children. Young children, given deliberately incomplete instructions for a card game, do not realize that instructions are inadequate until they play the game (Markman, 1979). Older children more readily realize the inconsistencies. Another metacognitive ability that often distinguishes developmentally different individuals, is the ability to assess the demands of the task. Older children relative to younger children realize that more study time for learning pictures results in better recall, and that paired associates that are opposites (good, bad) are easier to learn than are random pairs (ball, cigar) (as in Kreutzer, Leonard, & Flavell, 1975). Furthermore, older children relative to younger children know that active strategies of learning are more likely to result in superior learning than less active strategies (Kreutzer et al., 1975). Other areas identified by Brown (1978), in which children’s metacognitive deficiencies have caused problems, include predicting the outcome of strategy employment both before and after the use of strategies (Brown & Lawton, 1977), and monitoring the success of their attempts to learn (Brown & Barclay, 1976; Brown, Campione, & Barclay, 1978). Many of these metacognitive factors are also responsible for performance differences between learners who are classified as “normal” and those who are considered cognitively disadvantaged (e.g., learning disabled and mentally retarded). In fact, Brown and Barclay (1976) point out that the greatest problem with retardates may be their inability to use what they know.

Although research consistently indicates that metacognitive abilities distinguish cognitively disadvantaged learners and normal learners, as well as younger and older learners, the critical point is that metacognitive deficiencies are among the problems of most novices regardless of age. Novice chess players, for example (Chi, 1978), have metacognitive problems similar to those of young
card players (Markman, 1979). Similarly, novice x-ray technicians show inept scanning patterns (Thomas, 1968) like those of young children first learning to search a visual array (Mackworth & Bruner, 1970). It seems, then, that metacognitive abilities are related, at least in part, to the knowledge one brings to a situation. Therefore, it is not surprising that experts in particular academic areas, such as reading and mathematics, not only have more knowledge, but also display more effective metacognitive behaviors that permit them to apply their knowledge and cognitions.

Each of the cognitive factors discussed thus far contributes to successful or unsuccessful performance across a variety of academic domains. Because of this contribution, educators are interested in measuring these factors so that appropriate remediations can be made. Psychometricians must, therefore, draw upon the plethora of research in knowledge-rich areas that has been conducted by cognitive psychologists. In the following sections, research investigating the cognitive factors of knowledge, control processes, cognitive strategies, and metacognition within the academic domains of reading, writing, mathematics, and science are discussed. In addition, the implications of this research for measuring academic abilities are considered.

READING

The cognitive processes involved in reading have generally been divided into the two main components of decoding and comprehending (LaBerge, 1980; LaBerge & Samuels, 1974). In decoding, the reader matches the printed code to a known memory pattern and recodes the pattern into a string of sounds. In comprehending the reader imposes meaning upon the text. Automaticity in decoding is necessary for good reading ability because readers have a limited amount of processing resources they can allocate, and automatization of decoding frees up resources for comprehension. The relationship between decoding and comprehension has, in fact, been supported, because children instructed in decoding skills have subsequently improved their reading comprehension (Pflaum, Walberg, Karegianes, & Rasher, 1980).

The comprehension component of reading involves both literal and inferential comprehension. Literal comprehension requires the dual processes of lexical access and parsing. Put simply, lexical access is the process by which words are assigned meaning, and parsing is the process by which words are connected to form ideas. In inferential comprehension, the reader goes beyond the literal meaning of the text to integrate ideas, to summarize, and to elaborate upon the text with inferences and extrapolations.

Although it is sometimes useful to speak of decoding and comprehension as being separate components, they are actually interrelated and do not necessarily follow a “bottom-up” sequence going from the decoding of letters up through
literal and inferential comprehension. For example, Bartlett’s (1932) early work with story schemata, in which he observed reconstructive aspects of text recall, suggests that a “top-down” sequence is more parsimonious. Specifically, Bartlett believed that the reader’s declarative knowledge of the subject matter and of text structure guided the processes of decoding, lexical access, and parsing, such that meaning was constructed and not merely extracted by the reader. Further consideration is, therefore, given to the role that declarative knowledge plays in reading ability and to how such knowledge can be assessed.

Assessing Readers’ Declarative Knowledge

Several reading specialists have suggested that two sources of variance in reading ability are the degree of organization and elaboration of information in memory (Anderson & Reder, 1979; Frase, 1973; Meyer, 1977). Organization refers to the quality of hierarchical structures among categories and subsets of information in memory, whereas elaboration refers to the amount of links or pathways among memory nodes through which the activation of information can spread. Such characteristics of memory may be used to explain why, for example, skilled readers outperform less skilled readers on simple word matching tasks. For example, Ehri and Wilce (1983) compared young readers’ speed at reading familiar printed words such as “hat,” “boy,” or “car” with their speed at reading one digit numbers. Their results indicated that less skilled readers were slower at reading words than they were at reading digits; whereas, no such differences were observed among skilled readers. These findings suggest that differences in word matching ability may have more to do with semantic knowledge than processing speed. Other investigators have also found that, at younger ages, less skilled readers are slower at labeling letters and words (Frederiksen, 1981; Jackson & McClelland, 1979; Perfetti, Finger, & Hogaboam, 1978; Perfetti & Hogaboam, 1975; Vellutino, 1979). Notably, such differences between skilled and less skilled readers have not been observed beyond the fourth grade, however, which suggests that less skilled readers’ decoding processes may be slower because they have not developed well-organized and elaborate semantic knowledge structures.

Lack of well-organized and elaborate declarative knowledge may also help explain comprehension deficiencies. Bower, Black, and Turner (1979), for example, investigated how having knowledge about a topic facilitates inferential comprehension. Specifically, Bower et al. asked students to read a story about visiting the doctor and then had them recall what they had read. Because visiting the doctor was a familiar experience to most of the students, 20% of their recalls included information not found in the original story. They had filled in the details with information obtained from prior experience with visiting a doctor. The extent to which the readers’ recalls included such elaborations beyond the text may presumably reflect the amount of prior knowledge they had about visiting
the doctor. Such a technique could be used as a prereading test for measuring a student’s prior knowledge about a topic. Typically, teachers pretest students on their semantic knowledge of vocabulary that will be included in a reading assignment. Perhaps it is just as important that teachers pretest students on their experience with and knowledge about events that will be described within a passage. By having students read and recall short passages similar in content to the main reading assignment, teachers can then identify the extent to which readers’ recalls include elaborations beyond the text. Although this process might be tedious from a pedagogical standpoint, less skilled readers should nevertheless be pretested and be given additional declarative knowledge about a topic prior to reading.

Good readers not only have sufficient declarative knowledge about a particular topic, they also have knowledge about the structure of the specific discourse grammar (e.g., narrative, expository, or argumentative). Comprehension is, in fact, enhanced if readers have well-organized and elaborate discourse schemata in memory that serve to facilitate encoding and retrieval processes. Research has shown, for example, that skilled readers rely more upon the structure of a passage in developing a summary than do less skilled readers (Bartlett, 1978; Meyer, Brandt, & Bluth, 1980; Taylor, 1980). Specifically, Meyer et al. (1980) found that about three-fourths of good readers, one-half of average readers, and less than one-fourth of poor readers used text structure in their recall summaries. Those readers who did not use the text structure tended to simply list ideas from the text in a random fashion. These results suggest that good readers use text structure in recall because it is adaptable to their own schematic representation of text in memory. Essentially, then, meaningful interpretation of text requires well-structured and elaborate declarative knowledge about various discourse schemata.

Psychologists and educators are, of course, interested in how such discourse schemata may be assessed. The cognitive approach to assessing such schemata involves comparing a reader’s recall of a passage with a structural model that indicates the major elements of a text and the relationships among those elements. The structural model of the passage serves as a scoring template used to examine both the amount and the type of information recalled by the reader, thereby revealing differences between the text structure and the reader’s organization in recall. From this analysis, psychologists then infer the amount of discrepancy between the structural model and the reader’s schematic structure for a given discourse in memory.

Cognitive psychologists have developed different approaches for analyzing text structure (e.g., Frederiksen, 1975; Kintsch & van Dijk, 1978; Meyer, 1981), but have, as yet, failed to converge on a simple, widely accepted method. Meyer (1981) cites several reasons for this lack of agreement. First, interest in describing text structures has historically come from disciplines as diverse as rhetoric, folklore, linguistics, education, psychology, and artificial intelligence. Such
plurality in backgrounds makes it difficult for academicians to reach consensus. Second, because of these diverse disciplines, the purpose for which structural analyses were developed has varied from that of assessing recall of main ideas to that of assessing the integration of logical relationships. Finally, since reading is a “top-down” process, the structure of a text will be described differently by readers who possess different prior knowledge and experience. This will be particularly problematic when the inherent structure of a text is more implicitly than explicitly stated.

Despite these confounding variables that affect how text structure is analyzed, psychometrians should attempt to establish a standard analytical method not open to the subjective affects of prior knowledge. Perhaps computer programs would be beneficial for building objective structural models of various discourse types. The programmer could specify the type of discourse to be analyzed and the intent (e.g., to identify main ideas or to identify logical relationships), and then enter the specific passage into the system so that an objective structural analysis could be conducted. Subsequently, each reader’s recall could be entered into the system so that a “goodness of fit” comparison could be made between the computer-generated model and the reader’s recall protocol. Upon making the comparison, the computer could then specifically identify, for the reader, what discrepancies might exist between the organization of the structural model and the organization of the readers’ discourse schema in memory.

The importance of assessing declarative knowledge among readers has been well-established. An equally important cognitive factor that must be assessed is the control processes that operate within the information-processing system during reading. A discussion of these processes and how they may be measured is addressed in the following section.

Assessing Readers’ Control Processes

Individual differences in control processes may account for differences observed between skilled and less skilled readers in recoding ability. Recoding, which involves connecting a string of sounds, requires holding small bits of information in temporary storage until sufficient amounts have been received in order to apprehend meaning (Baddely, 1970; Conrad, 1972). Presumably, then, recoding might involve the control processes of attention, rehearsal, chunking, and the manipulation of information in working memory.

Research investigating speed of recoding reveals that less skilled readers are slower at starting to say pseudowords than are skilled readers (Frederiksen, 1981). Such deficits in recoding speed would be expected among less skilled readers, because their decoding processes have not yet become automatized. It is important to note, however, that differences in recoding ability have been observed to disappear by the third grade (Venezky & Johnson, 1973) and, consequently, one must again consider the role knowledge plays in performance of these tasks.
Additional research by Perfetti and Roth (1981) illustrates differences among skilled and less skilled readers in their ability to integrate sentences. Sentence integration requires the holding and the manipulating of information in working memory, because the reader must combine successive sentences in order to integrate ideas. Specifically, Perfetti and Roth (1981) asked students to listen to pairs of related sentences in which the last word in the second sentence was missing, and to then predict the last word in the second sentence. An auditory presentation was used in order to focus on the comprehension process and not on the decoding process. The results indicated that skilled readers, relative to less skilled readers, produced a greater number of appropriate sentence-ending words on moderate-constraint sentences. (These are sentences that can be ended with a moderate selection of possible words.) Apparently, then, less skilled readers were more likely to produce inappropriate words to complete the sentence because they were unable to hold the relevant information from the first sentence in working memory.

Related findings by Frederiksen (1981) indicate that less skilled readers' reading speed is slowed down when the second sentence in a pair contains a pronoun reference or an implicit reference to a noun phrase in the previous sentence. Presumably, the slower reading rate occurs because readers cannot hold an adequate amount of information in working memory and must, consequently, look back to the prior sentence to identify the noun. One explanation for why less skilled readers perform poorly on this task is that they have limited working-memory capacity. Findings that support this view include those that reveal deficits in short-term memory recall of digits (Corkin, 1974; Bakker, 1972; Jorm, 1977) and of word strings (Bauer, 1977; Torgeson & Goldman, 1977) among less skilled readers.

Despite the abundance of research supporting a capacity hypothesis, alternative hypotheses must be noted. One hypothesis is that individual differences in readers' working memory capacities are due more to differences in control processes than to differences in hardware. Such control processes as rehearsal and chunking are limited in their simultaneous application and, therefore, compete for the readers' attention. Most memory-span tasks require readers to simultaneously attend to incoming data while rehearsing information already temporarily stored in working memory. Therefore, it is the competition between these control processes, and not the capacity of working memory, that hinders performance on such tasks. A second hypothesis, tested by Daneman and Carpenter (1980), posits that less skilled readers do poorly on memory-span tasks because they do not perform some of the simpler literal comprehension processes (i.e., lexical access and parsing) as automatically as do skilled readers. Specifically, the authors devised a reading-span task whereby subjects read sentences aloud at their own pace and then attempted to recall the last item from each sentence. Results found the reading-span task to be a better predictor of verbal ability and of reading comprehension than was a conventional digit-span task. Daneman and Carpenter (1980) concluded, therefore, that less skilled readers...
performed poorly on the reading-span task not because they have a limited working-memory capacity, but because they have difficulty assigning meaning to words or putting words together.

Research investigating differences in control processes among readers has implications for the field of measurement. First, based on Daneman and Carpenter’s (1980) findings, it would be unwise to draw conclusions about the nature of a reader’s working memory capacity unless he or she does poorly across a wide variety of memory-span tasks. Second, performance deficits on memory-span tasks do not necessarily indicate a fixed capacity limitation in working memory. Developmental research suggests that individual differences observed on these tasks are due more to the use of control processes than to differences in architecture. Besides, the prospect of a fixed capacity limitation leaves little hope for the possibility of appropriate remediation. It is perhaps more reasonable for psychometrians to investigate methods for assessing the underlying deficiencies in control processes that characterize poor readers.

Assessing Readers’ Cognitive Strategies

Another factor to consider in assessing reading ability is the reader’s cognitive strategies. Cognitive strategies are methods for reaching some goal in an optimal way (van Dijk & Kintsch, 1983). They require conscious, controlled, cognitive representations that dominate the moves of an action sequence. The fact that cognitive strategies are conscious makes them amenable to measurement through verbal reports of individuals as they solve problems. Examples of what might be discovered from such reports are that individuals may break a problem down into subproblems, they may attempt to obtain more information to solve the problem, or they may return to previously solved states of a problem if an error is made.

Language strategies, which operate during reading, are unlike typical cognitive strategies because they are not consciously controlled. They occur almost automatically, particularly with continued practice. A number of actions are involved, for example, in reading that occur rather unconsciously, such as identifying letters, constructing words, analyzing syntactic structures, and understanding sentential and textual meanings. In spite of the effortless nature of these processes, van Dijk and Kintsch (1983) believe it is appropriate to speak of strategies that operate in discourse comprehension for the following reasons: (a) The language user is confronted with the task of understanding an action; (b) such an action has a well-defined goal (comprehension); (c) the solution occurs step-by-step, and may be broken down into subtasks; and (d) the solution is not always obvious, and therefore alternative routes may need to be taken (pp. 71-73). Essentially, the authors suggest that discourse comprehension is an instance of human problem solving and, therefore, necessarily requires the use of language strategies.

Language strategies are different from language rules (which more generally specify correct structures for phonology, morphology, or syntax) because they
are context-dependent. That is, rules describe proper structure for clauses, sentences, and paragraphs; strategies describe how these rules are employed within the context of the semantic analysis of a passage. Rules, then, have to do with syntax, whereas strategies have to do with semantics. The importance of language strategies for reading is that they efficiently apply abstract language rules in such a way that several levels of discourse can be processed simultaneously. The specific nature of these strategies, as described by van Dijk and Kintsch (1983), is now discussed.

**Propositional Strategies.** A proposition is a composite unit that includes a predicate and one or more arguments, where a predicate is defined as being a property or a relation, and an argument is defined as being a thing or a person. The unit "a boy" would not be considered a proposition, because it only contains an argument. The unit "a boy ran home" would be considered a proposition, however, because it contains both an argument and a predicate. Propositions are constructed by the reader based on the context of the passage and word meanings activated from semantic memory. Propositional strategies guide the reader in placing predicates and arguments into configurations, and in helping the reader make best guesses about the likely structure or meaning of incoming data. An example of a propositional strategy is assigning a noun or a pronoun as the subject of a proposition even before the rest of the clause has been analyzed. If such an assignment turns out to be wrong, then a second strategy would be to go back over the clause applying the rules of syntactic structure. These kinds of propositional strategies operate continually and facilitate automaticity in reading.

Research indicates that skilled readers are more proficient at using propositional strategies. Specifically, Frederiksen (1981) asked high school students to read sentences that had the last word missing and, after they had read each sentence, to press a stimulus that released the missing word. Students were then to pronounce the word as fast as they could. Frederiksen (1981) reasoned that if they were expecting the word, the students would pronounce it faster than if they were not expecting it. Two types of sentences were provided: those providing "weak context" and those providing "strong context." Again, the author reasoned that if good readers were more proficient at propositional strategies, they would benefit more from having the strong context than would the poor readers. Results, in fact, found that good readers did show a greater difference in reaction time between weak and strong context sentences than did poor readers. Frederiksen concluded, therefore, that skilled readers used propositional strategies to make several best guesses about word meaning possibilities, and were therefore prepared to pronounce any one of them.

**Local Coherence Strategies.** Local coherence strategies help to establish meaning among successive sentences. The assumption underlying local coherence strategies is that language users attempt to establish some coherent relation before they have fully processed a pair of sentences. They will do so by
relating fragments of the new sentence to the sentence previously processed. More specifically, local coherence will be established among sentences by searching for propositions that contain related facts or potential links, or by recognizing argument repetition that may be both explicitly and implicitly stated. Essentially, then, local coherence is strategic because relatedness among sentences must be established by the reader.

Research indicates that skilled readers are more proficient than less skilled readers at establishing local coherence. As has been previously described, Perfetti and Roth (1981) asked 8 to 10 year old students to listen to pairs of related sentences and to predict the last word in the second sentence. To perform well on this task, one must establish commonalities among the two sentences in order to make an accurate prediction. Again, Perfetti and Roth’s results indicated that skilled readers were more accurate than were less skilled readers with sentences that contain moderate-constraint sentences. These findings suggest that skilled readers are better able to integrate sentences efficiently because they employ strategies for integrating common propositions between sentences.

**Macrostrategies.** Macrostrategies operate at the level of macrostructures that describe the overall meaning or gist of a passage. Macrostructures are different from schema structures because the latter represent the form of a discourse grammar (e.g., a story schema contains information about plot, setting, resolution, and so forth). A macrostructure, on the other hand, is the global meaning inferred from a passage. In order to establish this global meaning, the reader must continually form best guesses about the main idea, even before he or she is finished reading. Macrostrategies, consequently, use propositions to form best guesses about a macrostructure that can, in turn, be used to understand subsequent sentences. This type of macrostrategy is described by van Dijk and Kintsch (1983) as semantic inference. Semantic inference is influenced by prior knowledge, by redundancy of propositions, and by macropropositions that are topical or thematic expressions that signal what the main idea is about. Such expressions often appear at the beginning or at the end of paragraphs, or may be signaled by larger print or by italics. Macropropositions that appear at the beginning of a section help the reader form hypotheses about the meaning of sentences to come; whereas those that appear at the end of a section serve to evaluate already established macrostructures.

The notion of discourse comprehension strategies seems useful if comprehension is considered to be a problem solving activity. The reader continually makes best guesses about how to solve the problem that concerns what the discourse is about. If, as van Dijk and Kintsch suggest, these strategies are not consciously controlled, how, then, can psychometricians devise tests to measure them? The traditional method of assessing cognitive strategies through verbal reports seems hardly valid in this case. The previously described tasks employed by Frederiksen (1981) and Perfetti and Roth (1981), however, seem useful for
assessing propositional and local coherence strategies, respectively. Such tasks are amenable to computer administration and scoring, and seem to come from a theoretical base closely aligned with the notion of language strategies. Steps should be taken to make these kinds of tasks readily available to those interested in assessing propositional and local coherence strategies. At the level of macrostrategies, cognitive psychologists have for years analyzed the free recalls of readers in order to assess the proportion of macropropositions they can remember. Educators should continue this type of testing so as to monitor whether readers can infer the main idea from a passage.

To validly and reliably assess discourse comprehension strategies, tests must require the examinee to actually connect propositions such that local and global coherence is established. The tasks cited above are a beginning, but innovative assessment devices must still be created. Recent research in the assessment of discourse production strategies (Benton & Kiewra, 1985), to be discussed within the writing section of this chapter, may provide insight into how comprehension strategies may be assessed.

Assessing Readers’ Metacognitive Processes

Investigations into the metcognitive processes of reading reveal differences between good and poor readers in their comprehension monitoring. Comprehension monitoring is a two-stage process of goal checking and remediating. In goal checking, the reader checks to see if he or she is achieving the goal of comprehension. Goals may vary according to whether one is reading for the purpose of skimming or for the purpose of obtaining a thorough understanding of a passage. During remediation, the reader looks back to previously processed discourse in order to pick up relevant information that was missed.

Differences have been observed between mature and less mature readers in their goal-checking strategies (Harris, Kruithos, Terwogt, & Visser, 1981). Specifically, Harris et al. (1981) asked third and sixth grade students to read stories, some of which contained an anomalous sentence relative to the title of the story (e.g., the sentence “He sees his hair getting shorter” within a story titled “John at the Dentist.”). Other stories containing the same sentence were more aptly titled “John at the Hairdresser’s.” Results found that reading speed was slower in the inappropriately titled stories for both grade levels, which suggests all students were cuing themselves that something was wrong with the anomalous sentence. Interestingly, however, 30% of the third graders could not identify the anomalous sentence, compared to only 11% of the sixth graders who could not. Apparently, then, students in both grade levels produced signals that their comprehension was faltering (because of a slower reading speed), but sixth graders were able to check the source of that signal. The authors contend, therefore, that mature readers are more adept at goal checking.
In order to investigate remediation skills, Garner and Reis (1981) examined a “lookback” strategy among students in the fourth through tenth grades. Specifically, students read passages containing successive paragraphs, each followed by three questions that required looking back to preceding paragraphs. Skilled readers, across all grade levels, looked back on an average of 30% of the questions as compared to less skilled readers who did so on only 9% of the questions. In addition, the six oldest readers looked back on 80% of the questions. Such findings suggest that mature readers employ remediation strategies in monitoring their comprehension.

The techniques used by Harris et al. (1981) and by Garner and Reis (1981) may prove useful as assessment devices for determining which students employ goal checking and remediation strategies during reading. Students should also be questioned individually about their use of metacognitive strategies in order to precisely identify the source of their deficiencies. Weinstein (1978), for example, has developed a questionnaire to assess readers’ strategies for elaborating upon a text. The questionnaire directs students to think about the purpose for their reading and to relate the passage to their own knowledge and experience. Similar questionnaires could be developed that assess the degree to which readers monitor their comprehension through goal checking and remediation strategies. Such questionnaires would presumably query readers about whether they understand the meaning of a passage and about their use of lookback strategies.

Summary

Research in cognitive psychology suggests that readers should be assessed with regard to their declarative knowledge, control processes, discourse comprehension strategies, and metacognitive processes. In assessing readers’ declarative knowledge, teachers must be encouraged to provide prereading assignments that test the reader’s knowledge about a given topic. In addition, in assessing the organization of declarative knowledge, psychologists need to establish a standard method for the structural analysis of text amenable to computer scoring. Second, before drawing conclusions about a reader’s working memory capacity limitations, he or she should be tested on a variety of memory-span tasks. Such limitations may actually have more to do with deficient control processes, however, than with deficient information-processing hardware. Third, discourse comprehension is established at the levels of propositional strategies, local coherence strategies, and macrostrategies. Propositional and local coherence strategies can be assessed with tests that require readers to integrate propositions both within and between sentences. Macrostrategies can be conveniently measured with free recalls of the main ideas contained within a passage. Finally, metacognitive processes, such as comprehension monitoring, can be assessed using tests that determine readers’ goal checking and remediation strategies. By focusing on these specific cognitive factors that operate during reading, educators can hopefully define specific skill deficits and provide precise interventions.
Within the academic domain of writing, John Hayes and Linda Flower (1980) of Carnegie-Mellon University have developed a model of writing formulated through direct analysis of writing processes. Their model proposes three interacting components within writing: (a) the task environment, (b) long-term memory, and (c) the writing processes. The task environment refers to the conditions surrounding the writing behaviors; that is, the writing assignment itself and the text generated thus far. The long-term memory component includes the writer’s declarative knowledge about the topic, the informational needs of the intended audience, and the overall plans that guide the writing processes. Within the third component of the model, Hayes and Flower (1980) describe three processes: planning, translating, and reviewing. Within the planning process, there are three subprocesses: goal setting, generating, and organizing. Goal setting refers to the purpose for writing and to the goals writers set for themselves. Generating involves accessing relevant information from long-term memory and the task environment to generate ideas for writing. Finally, in the organizing subprocess, the writer attempts to establish both cohesion and coherence in writing. Cohesion refers to the use of linguistic devices (e.g., pronouns, conjunctions, and implicit linguistic ties) that integrate related ideas. Coherence, on the other hand, refers to how well an entire passage fits together.

In translating, ideas (semantics) are transformed into external symbols (syntax). This is actually the direct opposite of decoding in the reading process, in which symbols are translated into ideas. Finally, in reviewing, the writer evaluates what has been written and makes revisions where needed. This process, therefore, involves the two subprocesses of evaluating and revising.

The components of the Hayes and Flower (1980) model are both iterative and interactive, because the writer continuously passes back and forth across these components during writing. Although the Hayes and Flower (1980) model is useful for identifying the various writing processes, it is, nonetheless, inadequate for investigating individual differences, because it fails to specify the cognitive factors that influence such processes. For this reason, a discussion of those cognitive factors follows, with particular attention given to how each may be assessed.

Assessing Writer’s Declarative Knowledge

Writing is perceived as an instance of information processing, because information must be retrieved from long-term memory to impose meaning on the specific writing task and to generate ideas for writing (Hayes & Flower, 1980). In order to write effectively then, writers must possess appropriate declarative knowledge in long-term memory. What kinds of knowledge contribute to expertise in writing ability? According to Perfetti and McCutchen (in press), relevant knowledge
Discourse schema knowledge refers to knowledge of discourse forms (Meyer, 1975; Stein & Glenn, 1979; Stein & Trabasso, 1981). More specifically, discourse schemata "include knowledge of the general structure and ordering of information within a given discourse, the typical qualitative nature of that information, and the kinds of linguistic ties that link that information into a coherent discourse" (Perfetti & McCutchen, in press, p. 42). Discourse schema knowledge would, for example, be important for someone trying to write a story. As we write a story, we construct the events of that story to correspond to our story schema which may include an introduction, a characterization, a resolution, and so forth. We do this because stories that are presented in a manner consistent with story schema structure are more comprehensible (Thorndyke, 1977). Knowledge of discourse structure seems essential, then, for expertise in writing, because such knowledge influences how prose is structured.

With regard to the Hayes and Flower (1980) writing model, the organizing and reviewing processes would seem to be most affected by such knowledge. Individual differences have been observed among writers, for example, in their ability to produce well organized text. Specifically, McCutchen and Perfetti (1982) compared text structures written by fourth and sixth graders. Students were asked to consider several constraints about a topic (e.g., the topic had to be about something both fun and dangerous), because the authors believed that the ability to simultaneously satisfy several constraints at once produces well-organized prose. Fourth graders tended to produce text with a listlike structure, considering one constraint at a time, whereas many sixth graders produced text with a zigzag structure that weaved back and forth across constraints. McCutchen and Perfetti also compared the students' writing with an ideally coherent text produced by the authors. They found that 60% of the sixth grade texts resembled the ideal structure, whereas only 44% of the fourth grade texts were so structured. Older students' essays were apparently better structured because of their more mature discourse schema structures. It seems, then, that skilled writers have acquired well-organized schema structures that assist them in organizing prose.

Similarly, discourse schema knowledge influences the reviewing process in writing. In fact, individual differences in the reviewing process are considered largely developmental in nature (Scardamalia & Bereiter, 1983), which underscores the importance of having adequate knowledge of discourse. Writers cannot, for example, effectively evaluate prose unless they have adequate schematic representations in memory with which to compare it. Writers with better organized and elaborate schemata for different discourse types will likely be more adept at establishing a goodness-of-fit between their prose and an ideal structure.
within memory. Consequently, they will be more proficient at both evaluating and revising their writing.

Stallard (1974) found that students differ in both the quality and the quantity of their revisions. Educators may want to use this diversity by employing cooperative writing methods that pair skilled with less skilled writers. Research has consistently demonstrated that cooperative learning facilitates academic achievement (Dansereau et al., 1979; Sharan, 1980; Slavin, 1980) as well as transfer to individual learning (McDonald, Larson, Dansereau, & Spurlin, in press). Cooperative learning has, in fact, been proposed as a useful instructional device in teaching writing (Gebhardt, 1980; Jacko, 1978). Educators have used this technique by creating peer response teams comprised of from two to five students who evaluate what each has written (Moore et al., 1986). In using peer response teams, however, teachers should encourage students to a) focus initially on what is done well, b) state negative reactions as questions, c) use either oral or written responses, and d) initially listen to all feedback before responding to criticism.

Cooperative learning is effective because it presumably provides the opportunity for observational learning and for immediate peer evaluation. Students who pair off and then write, exchange, and revise may assist each other in evaluating the quality of their schematic structures necessary for organizing and revising prose. This method seems more effective than the traditional pedagogical techniques of correcting errors and writing comments that require no academic response by the learner. Educators must realize, however, that additional findings suggest that teachers should still be involved in the evaluation of writing, because many students apply evaluative criteria significantly different from those of their instructors (Newkirk, 1984). Teachers who urge students to write solely for their peers may, therefore, reinforce writing that fails to meet the expectations of academic audiences.

Besides discourse schema knowledge, writers must possess lexical knowledge—knowledge of words and their meanings—as well as syntactic knowledge, along with procedures for coordinating that knowledge. Lexical and syntactic knowledge assist in the manipulation of ideas into their correct ordering within a sentence. The process in writing influenced by such knowledge would most likely be translating.

Writing blocks, which hinder automaticity in the translating process, may presumably occur if the writer lacks adequate lexical and syntactic knowledge. If the writer continually struggles to access a word or agonizes over concerns with grammatical structure, then the fluent translation of ideas will be blocked.

Effective writers apparently have methods for acquiring additional information so that translating is more automatic. They may read texts on writing style, or perhaps flip through a thesaurus if searching for the correct word. Whatever the method, one would expect that good writers have acquired the lexical and syntactic knowledge needed for facilitating automaticity in translating. Again, educators may assist writers in assessing their lexical and syntactic knowledge
through cooperative learning methods. Peer editing teams can help the writer monitor features of writing mechanics by providing feedback about spelling, punctuation, and word usage (Moore et al., 1986). Whereas peer response reacts to the writing as a whole, peer editing reacts to the specific structure of sentences.

In assessing lexical knowledge, educators may also want to pretest their students on vocabulary that would be relevant to a given topic. It may be of even greater value to test whether students can then generate sentences that contain certain vocabulary, because practice in using the words in writing may facilitate automaticity. In addition, tests that assess basic grammar, such as the Test of Standard Written English (TSWE) of the College Board (1983), should continue to be used for the purpose of assessing students’ knowledge of syntax, punctuation, and word usage.

In addition to knowledge of discourse and mechanics, expert writers must also have sufficient knowledge of specific topics. The extent of one’s knowledge about a particular topic would presumably influence the generating process in writing, because such knowledge contributes to the elaborateness and the relevance of ideas produced in writing (Voss, Vesonder, & Spilich, 1980). Voss et al., for example, asked college writers with equal verbal ability, but with varying degrees of knowledge about baseball, to write an account of one-half inning of baseball. Students’ written texts were then analyzed by categorizing propositions according to those dealing with game actions, auxiliary game actions, relevant nongame actions, and irrelevant nongame actions. Results indicated that writers with greater baseball knowledge generated a higher proportion of auxiliary game action propositions (e.g., elaborations about where a ball went when hit) than did those with limited baseball knowledge, whereas those with limited baseball knowledge generated a higher proportion of irrelevant nongame actions (e.g., propositions concerning the fans’ behaviors).

Recent advances have been made in measuring the influence of knowledge on the generating process based on structural analyses of students’ writing (Benton & Blohm, 1986). Specifically, Benton and Blohm contend that the generating process in writing can be measured by considering the extent to which writers elaborate upon their ideas with explanations and examples. Because ideas should be well organized, methods for measuring such elaborations in writing must be sensitive to the relationships between superordinate and subordinate ideas contained within a passage. This relationship can be broken down into three basic concepts that reflect both elaboration and hierarchical relationships: top-level, mid-level, and base-level ideas (Meyer, 1977). Specifically, ideas are top-level when they are related to an idea of central importance that relates several concepts together. Mid-level ideas are explanations, definitions, or descriptions that clarify the relationship directly stated or inferred in a top-level idea. Finally, base-level ideas provide specific details that exemplify a mid-level explanation or a top-level relationship. Consider the following example from a text generated
by a student who wrote on the topic "Wastefulness is a necessary part of the American way of life":

We, as Americans, are very wasteful (top-level idea). Each day millions of us get up out of bed and immediately begin being wasteful (mid-level idea). Soaps, powders, lotions, cosmetics, tissues, and other elements are consumed (five base-level ideas).

This type of structural analysis is useful for assessing generating in writing because it provides both a quantitative and a qualitative measure of elaboration. Within the preceding passage, for example, one can count a total of seven ideas. More importantly, however, there are five base-level ideas for each mid-level idea, and one mid-level idea for the single top-level idea. These types of measures indicate to what depth the writer elaborates upon mid- and top-level ideas.

Appropriate prior knowledge is an important prerequisite for good writing, and must therefore be assessed within the context of the various processing components of the writing model. Other aspects of the writer that should be considered are the control processes that operate within the translating component.

Assessing Writers’ Control Processes

Recent investigations within the domain of writing have identified individual differences in the information-processing system (Benton, Kraft, Glover, & Plake, 1984). These differences between good and poor writers (as defined from holistic impressions of writing samples) are reflected in the holding capacity of short-term memory, and the manipulation of information in working memory.

*Holding Capacity of Short-Term Memory.* In writing, as verbal information is transferred from long-term to short-term memory, it must be held there while translating processes are carried out. One must be able to hold letters together so that they may be put together to make a word, and words must be held together to make a clause. If the holding capacity of short-term memory is small, then presumably the process of language production will require more time and be less automatic.

In order to measure the holding capacity of short-term memory among good and poor writers, Benton et al. (1984) used a modified version of a task developed by Peterson and Peterson (1959). Subjects were presented with four consonants on a screen, one at a time, for .50 s followed by a distractor task of reading numbers from the screen for a variable amount of seconds. They were then asked to recall the four letters in their correct order. This task assessed holding capacity, because it required a person to hold information in short-term memory while concentrating on something else. Similarly, writing involves holding information in memory while deciding how to connect it to other information.
In a high school sample, good writers recalled significantly more letters in their correct order on this task than did poor writers, controlling for reading comprehension, reading speed, and scholastic achievement. These results suggest that the holding capacity of short-term memory is one factor that discriminates good from poor writers.

**Manipulation of Information in Working Memory.** Another important control process that is crucial for language production is the ability to perform rapid operations on information held in working memory. In order to write, a person needs to combine letters into words, words into clauses, and clauses into sentences.

When Benton et al. (1984) sought to assess writers’ abilities to manipulate information in working memory, they developed a letter reordering task. Good and poor writers were exposed to a sequence of five randomly selected consonants displayed for .50 s on a microcomputer display screen. They were instructed to hold the letters in working memory and to recall them in alphabetical order. Results found that good writers recalled significantly more letters in correct alphabetical order than did poor writers. These differences were observed in both high school and college samples when reading speed, reading comprehension and achievement were controlled.

The methods devised by Benton et al. (1984) are amenable to simple administration and scoring, particularly when using a microcomputer. As has been suggested with regard to working memory capacity among readers, however, writers should also be tested with several tests before conclusions are reached about any translating deficits. In addition, it would be wise to obtain post-hoc verbal protocols of writers that describe the cognitive strategies they may use while performing these tasks.

**Assessing Writers’ Cognitive Strategies**

As mentioned previously, Kintsch and van Dijk’s (1978) model of strategic discourse processing posits that comprehension strategies operate at several levels of discourse. Specifically, their model describes propositional strategies, which integrate words and clauses; local coherence strategies, which integrate successive sentences; and macrostrategies, which integrate macropropositions of the overall text. Although their model was originally developed for discourse comprehension analysis, van Dijk and Kintsch (1983) contend that the basic mappings between surface structure expressions and semantic representations are the same for both comprehension and production of prose, even though the reader and the writer are concerned with different aspects of strategic discourse. The model, therefore, seems appropriate for analyzing strategies employed during the writing process.
Recently, methods for assessing discourse strategies among good and poor writers have emerged. Specifically, good writers have been observed to perform more effectively on tests involving word reordering within scrambled sentences, sentence reordering within scrambled paragraphs, and paragraph assembly, which requires the ordering of sentences into multiple paragraphs (Benton & Kiewra, in press; Benton et al., 1984). Notably, good writers outperformed poor writers on these tests, in both high school and college samples, when reading comprehension, reading speed, general knowledge, verbal ability, and achievement were controlled. Each of these specific tests and their intended level of measurement is now discussed in greater detail.

**Word Reordering Test.** This test was designed to assess propositional strategies used in writing that integrate propositions within a sentence. Specifically, students are presented items that contain a scrambled sentence and are directed to unscramble each sentence as rapidly as possible and to write in the correct version of the sentence. Although there may be more than one correct ordering, students are told to provide only one response. An example of a scrambled sentence and its correct form appear below:

*Scrambled version:* Fight feels him with teases anyone must he boy who the.

*Correct version:* The boy feels he must fight with anyone who teases him.

The word reordering test presumably measures the writer’s abilities to detect clause boundaries and to integrate propositions. Specific propositional strategies that might be employed in this kind of test item include the following sentence parsing strategies.

1. Whenever you find a determiner, begin a new noun phrase (Clark & Clark, 1977). In the previous example, the writer who employs this strategy would begin a phrase with “The boy,” because that is the only logical noun-determiner combination.

2. Whenever you find a relative pronoun (that, which, who, whom), begin a new clause (Clark & Clark, 1977). Again, drawing upon the example given, the writer who uses this strategy would attempt to begin clauses with “who must,” “who feels,” or “who teases.”

It seems reasonable to assume, then, that such strategies for discourse production would be employed in the word reordering test, which requires writers to integrate scrambled propositions. Differences observed between good and poor writers on this test might, then, be attributed to differential use of propositional strategies.
Sentence Reordering Test. This test was devised in order to assess local coherence strategies. In this test, students are presented with a series of items, each containing a chronological paragraph whose order of sentences has been scrambled. Students are directed to order the sentences chronologically by placing the correct order number for events in the blank alongside each sentence. Although there may be more than one correct ordering, students are told only to provide one solution. An example of a scrambled paragraph with one possible solution appears below. An example of a scrambled paragraph with one possible solution appears below.

8 Subsequently, each day that Hugh did a better job of putting the food in his mouth instead of elsewhere, I rewarded him with peaches.
7 Hugh received no peaches.
1 Hugh had a great fondness for peaches.
3 I showed him the peaches he could expect and pointed out that he should put the food in his mouth, not on the floor.
5 I gave him the peaches.
2 I told him that he could have peaches for dessert if he did not mess his food up so much.
4 He did better, although liberal amounts of food still fell on the floor.
6 The next day Hugh was in an exuberant mood and scattered his vegetables far and wide.
9 He improved rapidly and was eventually willing to substitute other fruits for his reward.

The sentence reordering test measures local coherence strategies, because writers must connect successive sentences in a chronological fashion by searching for related propositions and potential links. In order to perform well on this test, writers must consider both previous sentences as well as the present sentence being processed. Using this logic, one can strategically determine that the sentence "Hugh had a great fondness for peaches" is the only one not dependent upon a previous idea. Consequently, this sentence is ordered first. Upon further investigation, the writer infers that the sentences number 2 and 3 above must necessarily be successive, because they contain the common ideas of "told him he could have peaches..." and "...showed him the peaches he could expect...;" as well as "...not mess up his food so much..." and "...not on the floor." One would expect, then, that writers who perform well on this test also efficiently employ local coherence strategies during writing.

Paragraph Assembly Test. The paragraph assembly test was designed to assess macrostrategies employed during writing. Specifically, students are presented with items containing one set of three scrambled paragraphs taken from an
essay originally generated by Bruning (1968). (In that original essay, each paragraph contained one topic sentence and three subordinate sentences.) Students are directed to correctly group the sentences into three, four-sentence paragraphs by placing a letter (A, B, or C) in the blank before each sentence. An example of a three paragraph set is presented below with the letters in the blanks representing an ordering into the three correct paragraphs.

B There are only 450 miles of paved roads in Mala.
C The only non-military high official in Mala is the premier.
A Aluminum mining has been especially productive for the northern region.
A The economy of Northern Mala is based on mining.
B There is only one telephone for every 15,000 inhabitants of Mala.
C The cabinet of the premier must be approved by a panel of military officers.
A About two-thirds of the work force in the north are involved in mining.
C The government of Mala can be classified as a military dictatorship.
B There are only 300 miles of railways in the entire country.
A Mining of all types provides about 80% of the income in the northern region.
B Mala's communication system would probably rank as the worst of all African nations.
C Whoever controls the Malan army controls the country of Mala.

The paragraph assembly test presumably draws upon the macrostrategy of semantic inference, because writers must infer three basic topics from twelve sentences. Specifically, the writer must make subtle differentiations among the sentences because all twelve sentences deal with the same basic topic—the mythical nation of Mala. Because semantic inference is influenced by prior knowledge, different readers will derive different inferences from the same text. Knowing this, writers must attempt to constrain this kind of personal variation in interpretation through textual signaling of the main theme or topic throughout the passage, such that the sentences within the text share similar ideas. It is imperative, then, that good writers be able to differentiate between closely related concepts, so that only similar ideas are grouped together in a paragraph. The paragraph assembly test attempts to assess this ability, because writers must impose meaning upon groups of sentences that do share similar ideas. Macrostrategies thus come into play because the writer forms best guesses about the theme that connects a group of sentences, and then reads further to evaluate whether such hypothesized macropropositions are correct.

Recent research by Benton and Kiewra (in press) has investigated the concurrent validity of the word reordering, sentence reordering, and paragraph assem-
bly tests with measures of writing ability. Results have indicated that these tests are significantly correlated with holistic impressions of writing samples.

Admittedly, however, it is difficult to actually construct tests that uniquely assess these strategies. Within the sentence reordering test, for example, propositional strategies are involved in reading each sentence. Similarly, within the paragraph assembly test, both propositional and local coherence strategies are involved in finding common propositions between sentences. According to the Kintsch and van Dijk model, however, one can not devise a test that uniquely measures macrostrategies or local coherence strategies, because such strategies are interrelated.

These language strategies are apparently involved in the organizing component of the planning process. Speculatively, writers use propositional strategies to organize words and clauses, local coherence strategies to organize sentences, and macrostrategies to organize paragraphs. Psychometricians should, perhaps, use measures such as the word reordering, sentence reordering, and paragraph assembly tests to assess these strategies that facilitate well-organized prose. Such measures can be easily administered and scored, and may have more validity than verbal reports, because language strategies are not consciously controlled. Processes that are more consciously controlled, referred to as metacognitive processes, are discussed in the next section.

Assessing Writers’ Metacognitive Processes

Although research investigating metacognitive processes involved in writing is still in its infancy, methods for assessing such processes, as they influence the translating and reviewing components of the writing model, have emerged.

One writing process influenced by metacognition is translating. As has been previously mentioned, automaticity in translating may be hindered by blocking, a common psychological phenomenon that hinders effective communication in any setting. Cognitive therapists (e.g., Arnkoff & Glass, 1982; Beck, Rush, Shaw, & Emery, 1979) have identified cognitive components of blocking (e.g., distorted thinking, automatic thoughts, inferences, and assumptions that appear in “self-talk” of patients) that inhibit effective therapeutic intervention. Generally, these therapists help patients recognize and record faulty cognitions and teach them new “self-talk” statements that are more adaptive.

Similar progress has been made in the investigation of self-talk during writing (Boice, 1985). Boice has identified seven components of faulty metacognitive processing that impede effective translating (1985, pp. 97-98): (a) self-talk about the aversiveness of writing; (b) self-talk that justifies avoiding or delaying writing; (c) self-talk that reflects burnout, anxiety, panic, or groundless worries; (d) self-talk concerned with achieving more in less time or of unnecessary deadlines;
(e) self-talk indicating internal criticism that allows no mistakes or imperfections; (f) self-talk about fears of rejection; and (g) self-talk about maladaptive strategies for writing (e.g., favoring a single draft over revisions).

It seems apparent that these kinds of maladaptive thoughts will prevent effective communication in writing. Perhaps effective writers are characterized not just by specific writing skills, but also by “healthy” metacognition. Educators may do well, then, to interview students who find writing aversive in order to identify faulty cognitions that impede the translating process.

Metacognitive strategies also influence the reviewing process in writing. During the early elementary school years, writers first learn to evaluate whether something is wrong with their prose, but they may not be capable of revising it until the later elementary school years (Scardamalia & Bereiter, 1983). Scardamalia and Bereiter, for example, asked elementary students to evaluate and revise each sentence as they wrote an essay. The quality of both the evaluating and the revising was then judged by expert adult writers. Results found that 85% of the time fourth grade students could recognize that something was wrong with their writing, but 70% of the time they could not remEDIATE problems they had identified. Older students, on the other hand, were consistently successful at revising a problem they had recognized. These findings suggest that less skilled writers have the necessary knowledge for evaluating their writing, but often fail to take the required steps to revise. What can educators do to facilitate those steps? Again, the peer editing and peer response teams described earlier may be useful, because students learn to edit and revise errors they would not normally identify themselves.

Research into the metacognitive strategies employed during writing have implications for the field of measurement. Most multiple-choice tests of writing ability assess the writer’s skills at recognizing errors in sentence structure, punctuation, and syntax. If, as the previously cited findings indicate, most writers can recognize problems in their writing, but may fail to remediate them, then the key variable in the reviewing process—revising—is not actually being tested. Multiple-choice tests may, then, be of questionable validity. In fact, the Conference on College Composition and Communication, declared in the 1970s that multiple-choice measures of writing were narrowly focused and provided gross distortions of writing competence (Troyka, 1982). These kinds of tests, nonetheless, continue to be widely used.

Recently, however, writing samples have also been used with greater frequency for the purpose of measuring writing ability. Although writing samples appear to be more valid measures, they are too frequently first drafts, because the writer is only allowed a set time period in which to write. Consequently, the essay is then handed in, with no opportunity for revisions. Both multiple-choice tests and writing samples, then, fail to assess the writer’s ability to revise, which is an essential component of writing. In fact, Stallard (1974) has observed that one
important distinction between skilled and less skilled writers, which is often overlooked, is that skilled writers make more revisions.

Psychometricians must direct their efforts, therefore, toward assessing how writers make revisions in their writing. One method for doing so would be to employ short-answer items that require writers to rewrite sentences or even paragraphs that contain flaws. The drawbacks of such items are many in terms of scoring; but unless writing assessment moves in this direction, tests of writing ability will remain of questionable validity.

Summary

Research into the cognitive factors that influence writing has several implications for the field of measurement. First, educators must be encouraged to assess prior knowledge as it affects the generating, translating, organizing, and reviewing components of the writing model. Specifically, students should be tested concerning their prior knowledge of the writing topic, their vocabulary, and their knowledge of basic grammar before writing actually begins. In this way, instructional interventions can be made that will facilitate the generating and the translating processes in writing. In addition, teachers should be encouraged to write an ideally coherent essay with which to evaluate the structure of students’ essays. Structural models of both the teacher’s and the students’ essays can then be made in order to compare the organizational structure within each. Students’ essays can also be assessed using an analysis by Benton and Blohm (1986) that counts the number of top-, mid-, and base-, level ideas within an essay. This type of scoring system indicates both the quantity and the quality of elaboration in writing. Finally, students can assess their own knowledge of discourse structure by exchanging their essays with one another and by receiving immediate feedback about their own skills at evaluating and revising their prose.

With regard to control processes, writers should be assessed on several kinds of memory-span tests before being diagnosed as having a limited working memory capacity. In addition diagnosticians who use such tests should be cautioned against interpreting results as being indicative of a fixed capacity limitation. Finally, post-hoc verbal reports might be employed to ascertain the control processes writers use on these tests.

Strategies for discourse comprehension have been successfully measured by tests designed specifically to assess propositional strategies, local coherence strategies, and macrostrategies. Specifically, the word reordering, sentence reordering, and paragraph assembly tests might be used along with writing samples or multiple-choice tests to measure local and global coherence in writing ability.

Finally, research into the metacognitive processes involved in writing suggests that tests must be designed to assess writers’ revising skills, because most students can recognize errors in their writing, but may fail to revise them.
Unfortunately, most multiple-choice tests of writing ability assess evaluating but not revising.

The study of cognitive factors in academic abilities has been quite extensive with regard to reading and writing. Although less work has been done in the areas of mathematics and science, individual differences in cognitive factors are, nonetheless, apparent in those domains as well. The exact nature of those differences and how they can be measured is now discussed.

**MATHEMATICS**

Generally, educators separate mathematical ability into two broad components: (a) computation, which involves the application of algorithms and rules for carrying out mathematical operations; and (b) conceptualization, which requires problem representation and the application of heuristics and problem solving strategies. Although Briars (1983) has suggested that cognitive factors, especially prior math achievement, are the best predictors of math computation and conceptualization, the literature is sparse with regard to how such factors may be measured. A discussion follows, nonetheless, of how prior knowledge, control processes, cognitive strategies, and metacognitive processes in mathematics performance can be assessed.

**Assessing Math Students’ Prior Knowledge**

Several investigators have observed that skilled math students organize their declarative knowledge differently than do less skilled math students (Chartoff, 1977; Hinsley, Hayes, & Simon, 1977; Krutetskii, 1976; Silver, 1979). Silver (1979), for example, found that skilled math students organize knowledge according to categories of solution methods, whereas less skilled math students organize knowledge according to categories of problem contents. Specifically, Silver asked seventh-grade students to categorize 16 word problems that varied in both their content and in their solution methods. Students subsequently solved the same 16 problems and, based on their performance, were categorized into good, average, and poor math problem solvers. The author found that good math problem solvers grouped problems together on the basis of solution methods. Poor math problem solvers, on the other hand, had grouped problems together on the basis of the problem content. In addition, Silver found that the ability to categorize problems on the basis of solution similarities was strongly correlated with standardized measures of mathematics ability.

Chartoff (1977) employed a similar procedure with secondary and postsecondary level students. The students in that study were asked to rate the similarity of algebra word problems. The results again showed that the most important dimension for categorizing algebra problems was how they were solved.
These findings suggest that proficient math students are able to grasp the formal structure of math problems, and that they possess schemata for various types of problem solutions. What this implies is that good math students perceive a problem structure prior to its solution. Apparently, they establish a goodness of fit between their schemata for solution methods and a given math problem before solving it.

The knowledge discussed thus far has most to do with problem representation, an aspect of the conceptualization component of mathematics. Besides possessing well-organized structures in memory for problem representation, however, skilled math students likely possess adequate procedural knowledge for various mathematical computations. R. M. Gagné and Paradise (1961), for example, advocated the importance of assessing prerequisite procedural knowledge (intellectual skills) required for solving linear equations. Such preassessment, obtained by employing rational task analysis, can uncover lower-order skills that must be mastered prior to performing higher-order skills. By specifying which prerequisite skills must be mastered, psychologists have discovered that low achieving math students commit errors because they lack prerequisite procedural knowledge (e.g., knowledge of how to find a common denominator or of how to simplify fractions). Brown and Burton (1978), who refer to lack of procedural knowledge as “bugs,” developed the BUGGY program specifically to assess errors made in subtraction. They translated subtraction procedures into a computer program capable of 100% accuracy in computation. Changes were then made in the program to mimic students’ errors in order to see if the same patterns of errors emerged from the computer. These kinds of analyses are valuable because they provide a map for more specific diagnosis of students’ errors. In using the methods of rational task analysis and computer programming, then, educators can diagnose specific deficient skills by testing less skilled math students at each prerequisite step. In this way, specific remediations can be made and the student can advance to the next skill level.

Essentially, then, educators must assess two types of knowledge in mathematics performance: declarative knowledge of problem representation, and procedural knowledge of mathematical computation. These two types of knowledge presumably require different kinds of tests. Specifically, knowledge of problem representation requires discriminatory tests such that students “may progress almost without limit in such functions as understanding, critical thinking, appreciation, and originality” (Anastasi, 1982, pp. 97–98). Tests assessing conceptual knowledge of mathematics need to allow for individual differences among students’ achievement, because complete mastery of this domain is not possible.

Tests that assess procedural knowledge of mathematical computation, however, require mastery tests in order to determine whether or not the examinee has acquired the prerequisite skills. Ultimately, the purpose of such tests is to deter-
mine whether or not more instructional time is needed for each student, and not
to determine individual differences (Hanna, 1981).

Assessing Math Students’ Control Processes

Another cognitive factor that affects mathematical computational ability that
must be considered is that of control processes. Several researchers have studied
the relationship between mathematical computational ability and performance on
memory-span tasks. Speigel and Bryant (1978), for example, examined the
relationship between speed of information processing, intelligence, and math
achievement in 94 sixth grade students. They used a sentence-picture comparison
task, a pictorial similarities-and-differences task, and a matrix analysis task,
similar to the Raven’s Progressive Matrices. They found processing speed to be
correlated \( -.40 \) with math computation scores. With intelligence controlled, how­
ever, the relationship between these measures was almost negligible. Such findings
suggest, then, that processing speed does not contribute uniquely to math
achievement, and is probably more related to general intelligence.

Webster (1979) examined differences in memory-span between mathе­
mathematically proficient students (those performing at or above grade level on the
WRAT arithmetic subtest) and a group of “mathematically disabled” students
(those performing 2 or more years below grade level). Subjects were tested using
memory span for seven digits and for strings of seven nonrhyming consonants
that were presented both aurally and visually at one second exposure and one
second intervals. Results indicated that the mathematically disabled group had
significantly lower memory-span scores than did the mathematically proficient.

Overall, these findings suggest that memory-span performance is related to
mathematical computation skills, but that speed of information processing is not.
As has already been suggested within the domains of reading and writing, multiple assessments of memory-span performance should be made before conclu­
sions are reached about deficits in a student’s control processes or information­
processing machinery. Again, differences in memory span do not necessarily imply that proficient math students have larger working memory capacities. It is
more likely that they are able to allocate attentional capacity economically such
that they can simultaneously hold and manipulate information in working memo­
ry. It has been suggested that such attentional capacity develops concomittantly
with proficiency in mathematics achievement, and is not necessarily a precursor
to good computational skills (Briars, 1983).

Another promising area of test development may be that of assessing cog­
nitive strategies related to mathematics achievement. Because of the limited
amount of studies in this area, however, the assessment of cognitive strategies
and metacognitive processes in mathematics are addressed in one section.
Assessing Math Students’ Cognitive Strategies and Metacognitive Processes

The strategies used by proficient math students generally may be broken down into three types: (a) heuristics, (b) awareness of problem-solving processes, and (c) belief systems (Briars, 1983). Heuristics would fall into the general category of cognitive strategies as described in this chapter, whereas awareness of problem-solving processes and belief systems would be considered metacognitive processes.

Heuristics are more or less rigid operating routines that serve to narrow the potential behavioral alternatives considered by a student when confronted with a problem. These types of strategies are independent of content, because they may be applied within the context of any problem. Examples of heuristics used in solving math problems include drawing a diagram or thinking of a similar problem solved previously. These types of heuristics enable math students to impose a more meaningful representation upon a problem.

Awareness of one’s problem-solving processes, a metacognitive skill in mathematics, can help math students recall and execute appropriate routines. This awareness is beneficial in making two types of decisions: (a) tactical decisions about selecting the appropriate method, and (b) strategic decisions, which involve decisions about how one allocates time (Schoenfeld, 1979). Poor strategic decisions may be the most costly among less skilled math students. Schoenfeld (1979) has observed, for example, that less skilled math students do not make good strategic decisions, because they do not monitor their progress toward a solution (e.g., they spend 10 minutes calculating the area of a triangle without considering what that will contribute to the final solution.)

Finally, belief systems and expectations about math can affect math performance. Examples of faulty beliefs that impede math performance include the following (Lester & Garofalo, 1982; Silver, 1981):

1. The difficulty of a problem depends on the size of the numbers.
2. Problems require the application of only one math principle for their solution.
3. Key words appear only in the last sentence of a problem.
4. There is only one correct way to solve a problem.
5. Problems should take only a few minutes to solve.

It is easy to understand how beliefs such as these can impede success in mathematics.

Cognitive strategies and metacognitive processes in mathematics performance have been neglected by psychometricians, even though they appear to be important factors in such performance. Some heuristics could be easily assessed by
having students hand in their worksheets, which show how they solved the problems, along with their test. Other types of heuristics will presumably have to be measured through verbal reports of math students as they solve problems. Students should also be asked to keep a time log of their solution steps, so that the strategic decision making of time allocation can be assessed. Finally, beliefs about mathematics can be assessed with questionnaires that ask students to evaluate the truth or falsity of statements such as those listed above.

Summary

It seems that psychometricians have done a good job of designing mastery tests to measure procedural knowledge in mathematics computation. They may need to do more, however, in developing discriminatory tests that measure other cognitive factors related to math achievement (e.g., organization of declarative knowledge, heuristics, and metacognitive processes). The methods used by Silver (1979) and Chartoff (1977) for assessing students' declarative knowledge, for example, should be adapted for testing students' organizational structure for math concepts in memory. Asking students to record their solution strategies on paper, along with an approximate time log, may also go far in identifying those students who make poor strategical decisions in solving math problems. Again, the need for assessing these kinds of cognitive factors in mathematics may encourage psychometricians to move away from the mastery model and toward the realm of discriminatory tests.

SCIENCE

As in mathematics, the domain of science is sparse with regard to studies investigating individual differences in cognitive factors. Two factors that seem most relevant to studying individual differences within science, however, are prior knowledge and cognitive strategies.

Assessing Science Students' Prior Knowledge

The assessment of prior knowledge in science must access both procedural and declarative knowledge. In measuring procedural knowledge, rational task analysis and the mastery model have been applied in science as they have in mathematics. Okey and Gagné (1970), for example, performed a task analysis of the prerequisites needed to solve solubility-product problems in chemistry. (The solubility problem, common in chemistry classes, concerns the question of whether or not a solid matter will form when two chemicals are mixed together.) Based on a hierarchy of prerequisite skills needed for solving solubility prob-
lems, the authors instructed and tested the students at each level of those skills. They then tested the students on solubility type problems and found that performance increased as knowledge of prerequisite procedures increased. The authors concluded, therefore, that success in science problem solving is associated with knowledge of prerequisite procedures.

In addition to knowledge of prerequisite procedures, the organization and the content of one’s declarative knowledge influence science problem solving. Specifically, Chi, Feltovich, and Glaser (1981) gave Ph.D. physicists and students, who had had one course in physics, twenty category labels for describing physics problems. In response to these labels, the subjects were to tell all they could about problems subsumed within the label and how they might be solved. Based on these responses, the authors constructed a network of declarative knowledge to reflect each subject’s organizational structure for scientific declarative knowledge in memory. Results found that the experts’ memory structures contained more physics principles and a more hierarchical organization than did the novices’. As in the previously discussed domains of reading, writing, and mathematics, then, expert science problem solvers have more elaborate and better organized memory structures than do novices.

The observed differences between experts' and novices' prior knowledge have implications for the measurement of scientific abilities similar to those in mathematics assessment. As in mathematics, educators should test their students on prerequisite procedural knowledge required for solving science problems. In so doing, they will be able to diagnose specific skill deficits that, with proper instruction, can be remediated. In addition, structural models of students’ declarative knowledge can be constructed and compared to well-organized structures of scientific principles (e.g., as in the structure contained within science textbooks). Analyses of students’ memory structures could presumably reveal the “missing links” that need to be learned for better understanding of specific scientific principles.

Assessing Science Students’ Cognitive Strategies

Of equal importance to prior knowledge are cognitive strategies, which aid in both understanding and in solving science problems. Specifically, individual differences have been observed in the strategies used for understanding a problem, and in the types of problem-solution paths generated.

Within the realm of the social sciences, for example, Voss, Tyler, and Yengo (1983) discovered differences between novices and experts in how they represented a problem. These authors focused on one type of social science problem: an undesirable state of affairs that requires improvement. Specifically, they compared the thinking-aloud protocols of political scientists, whose specialty was Soviet politics, with those of college students taking a Soviet political science course. Each subject was asked to assume the role of the Soviet Ministry
of Agriculture and to consider the problem of how to improve productivity after experiencing low agricultural productivity during the previous 5 years. Voss et al. found that 24% of the experts’ protocol statements were devoted to defining the problem, whereas almost none of the novices’ statements were so devoted. More specifically, experts began by defining the constraints of the problem (e.g., Soviet ideology, soil conditions, and so forth), whereas novices simply began by listing possible solutions. What these findings suggest is that experts seek a deeper understanding of a problem before attempting to generate solutions.

Further findings suggest that individuals differ not only in how they represent a problem, but also in their problem-solution paths. It has been observed, for example, that novices engage in solution searching by attempting several paths toward reaching the goal of the correct solution; experts, on the other hand, follow one solution path (Larkin, McDermott, Simon, & Simon, 1980). These authors asked novices and expert physicists to think aloud while solving a problem for determining velocity. The experts used a “working forward” solution path, because they began with that which was known and proceeded step-by-step to the solution. Novices, on the other hand, used a “working backward” approach, because they began with the goal (solving for velocity) and tried to solve for it immediately before completing prerequisites steps. Based on this kind of finding, Gagné (1985) has drawn an analogy between being lost in a forest and solving science problems that captures the essence of the expert-novice distinction:

If one is lost in a forest, one is better off determining the direction (N,S,E,W) of one’s goal and limiting one’s search for a path to this direction than wandering around at random. The difference between novices and experts is that experts are not lost; they know a path that leads to the goal and follow it. (p. 282)

Even with regard to problem-solving strategies in science, then, one cannot underestimate the importance of prior knowledge, because one must have some knowledge of the constraints involved before representing a problem adequately. Similarly, in order to select the appropriate solution path, one must also have prior knowledge of what will or will not work. An important distinction between the novice and the expert, however, may be that novices do not bother to acquire the needed prior knowledge before generating possible solutions.

Summary

As in mathematics, psychometricians have probably been successful at assessing procedural knowledge in science problem solving using rational task analysis and mastery tests. More can be done, however, in constructing discriminatory tests that measure the organization of declarative knowledge and the application of problem solving strategies. Tests that require students to categorize scientific
problems (as in Chi et al., 1981) may be useful for assessing the organization of structures for science concepts in memory. In addition, verbal protocols could help to reveal the cognitive strategies employed by students as they solve science problems. Such protocols would presumably vary within as well as between students, however, depending upon their prior knowledge of the content area.

CONCLUSION

Research in cognitive psychology has identified underlying cognitive factors that contribute to intellectual performance across academic domains. If psychometricians are to follow the lead of cognitive psychologists, they must devise tests that measure domain-specific prior knowledge, control processes, cognitive strategies, and metacognitive processes. In so doing, measurement will become more valid and precise, and instructional treatments will presumably become more effective.

A major goal of measurement should be to assess the declarative knowledge structures that influence performance in various academic tasks, as “the nature and power of students’ organized structure of knowledge is a key aspect of educational achievement, because it either facilitates or hinders what he or she can do in a subject area” (Messick, 1984, p. 217). In fact, Glaser (1984) criticizes much of the research favoring the importance of control processes, because it was conducted in “knowledge-lean” domains (p. 94). Glaser further argues that it is the interaction of control processes with domain-specific knowledge that produces expertise. Psychometricians should, therefore, develop tests that assess students’ knowledge structures within each academic domain, using the structural model approach of cognitive psychologists. In this way, network or tree structures can be constructed that indicate symbolically the elements of a person’s knowledge and the relationships among those elements. These structures can then be compared to ideally organized knowledge structures so that teachers can identify students’ schemata that need to be restructured or elaborated.

In addition to declarative knowledge, learners should be assessed with regard to their procedural knowledge. Educators have actually been quite proficient in measuring procedural knowledge, as in the domains of mathematics, and science by using rational task analysis. They must, however, make greater use of computer programs (e.g., BUGGY), which list step-by-step actions to be taken, and of flowcharts, which present processes and decisions to be made, in order to specifically pinpoint students’ procedural errors. Ultimately, then, teachers will become more successful in identifying and in remediating specific procedural deficits.

In assessing control processes, memory-span tests, similar to those described in previous sections of this chapter, should be used to measure attentional capacity, rehearsal, chunking, and manipulation of information in working memory.
Hopefully, then, students who lack automaticity in decoding or in translating, which are affected by such processes, can be identified for remedial instruction.

Cognitive strategies should also be assessed across all academic areas. Such strategies, which are means for achieving a designated goal, may or may not be consciously controlled. Consciously controlled strategies (e.g., representing the problem, means–ends analysis, and so forth), which operate particularly within math and science, can be assessed through verbal reports of students as they solve problems. These kinds of data should continue to be collected and analyzed, and then programmed into computers to simulate human problem solving. Strategies that are not consciously controlled (e.g., language strategies within reading and writing), however, require alternative methods of assessment. They may be assessed with tests that tap such strategies, as in the word reordering, sentence reordering, and paragraph assembly tests used to measure discourse production strategies in writing.

Finally, more emphasis should be placed on the measurement of metacognitive processes. These processes are important because they facilitate the self-checking and the remediating processes involved in all academic areas, and because they orchestrate and control the previously discussed cognitive factors. Innovative assessment methods, which have been described within this chapter, should be employed to develop tests that measure metacognition within each academic domain.

The notion of developing tests, based on known cognitive factors that differentiate between experts and novices, seems in contrast to the more traditional approach of testing in which items are based on vague curriculum objectives, which may or may not be related to expertise (Messick, 1984). Tests that assess domain-specific cognitive skills seemed reasonable even to Cronbach (1957), however, nearly 30 years ago, when he wrote:

In dividing pupils between college preparatory and non-college studies... a general intelligence test is probably the wrong thing to use. This test, being general, predicts success in all subjects, therefore tends to have little interaction with treatment, and if so is not the best guide to differential treatment. We require a measure of aptitude which predicts who will learn better from one curriculum than from the other. (pp. 680–681)

Essentially, then, the cognitive approach to the assessment of academic abilities is closely related to the long-standing specificity doctrine first proposed by correlational psychologists. Those who hold this view, and those who adhere to the current cognitive approach, contend that individual differences consist of nothing more than differences in specific knowledge and skills acquired through learning. Cognitive psychology’s greatest contribution to the assessment of academic abilities will, therefore, probably be the notion that academic deficits have more to do with lack of domain-specific knowledge and skills than with lack of general intelligence.
REFERENCES


estimating test readiness: Generalization from list learning to prose recall. Unpublished manuscript, University of Illinois.


