

Rapid Assessment of Learners' Proficiency: A cognitive load approach

Slava Kalyuga*

University of New South Wales, Sydney, Australia

To tailor instruction to levels of learner proficiency in a domain in adaptive learning environments, rapid on-line methods of assessing learner knowledge and skills are required. Cognitive studies indicate that major factors in expert performance are organised structures in the person's knowledge base. Measuring different levels of acquisition of such knowledge structures should be a major purpose of cognitive diagnostic assessment. Traditional tests may not provide reliable evidence for diagnostic purposes and are not suitable for dynamic learning environments. This paper discusses an alternative diagnostic approach based on monitoring immediate traces of students' knowledge structures in working memory. Results of an experimental study applying this method to assessment of user reading skills are described. The study demonstrated a sufficiently high level of correlation (.66) between obtained measures and traditional test scores, and a substantial increase in reliability, with testing time reduced by a factor of 3.8.

A series of studies demonstrated that instructional techniques and procedures that are relatively effective for novices may become ineffective for more proficient learners (the expertise reversal effect; see Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga, Chandler, & Sweller, 2001). An instructional implication of this effect is that learners' proficiency levels should be accurately and continuously monitored in order to tailor instructional designs appropriately. Traditional tests are usually not suitable for this purpose. Apart from being slow and time-consuming, they do not always provide reliable evidence for diagnostic purposes.

For example, observing a student's correct answers to a series of simple equations (e.g., $5x = -4$) in a multiple-choice test would not tell us how those equations were actually solved. The student could have solved them by using a novice-like random search, or a trial-and-error method, systematically substituting different values for x until the right one is found. Alternatively, she or he could have solved these equations

*School of Education, University of New South Wales, Sydney, NSW 2052, Australia. Email: s.kalyuga@unsw.edu.au

by using an expert-like knowledge-based application of appropriate schematic solution procedures. In the latter case, different levels of procedural knowledge could have been applied: slow low-level fine-grained step-by-step procedures (e.g., dividing both sides of the equation by 5, and then cancelling the same numbers in the numerator and denominator on the left side of the equation), or higher-level automated procedures with final answers obtained immediately.

For another example, consider the assessment of reading skills. Traditional methods of testing reading comprehension are usually based on asking students to read sentences or segments of text and then answer multiple-choice questions about those sentences or texts. These methods are limited in their ability to establish different levels of reading skills and diagnose students' reading problems. A student's correct answers to a series of multiple-choice questions about a concurrently displayed textual passage do not tell us in what cognitive processes the reader was actually involved before selecting her or his answer, and by what method those answers were actually attained. The student could have repeatedly searched the text for key question words. Poor readers often read the questions first and then search the textual passages for answers to those questions. Alternatively, she or he could have constructed coherent mental representations of the text (a more proficient reader's approach). Thus, traditional reading test formats are not guaranteed to measure students' cognitive structures constructed during reading (Magliano & Millis, 2003). Even if the test formats prevent students from using back-search strategies, these tests do not always provide answers to some important diagnostic questions (e.g., is a reader capable of handling relatively simple sentences but are there problems with structurally more complex sentences, and at what level of complexity do the difficulties begin to appear?).

Enhancing the diagnostic potential of educational assessment using our current knowledge of cognitive processes and structures involved in learning and competent performance is a clear trend in the field. For example, a recent major report on educational assessment commissioned by the USA National Research Council (Pellegrino, Chudowsky, & Glaser, 2001) emphasised that a model of cognition should be the cornerstone of an assessment design process directed toward evaluating students' schematic knowledge structures. Cognitive science has accumulated a vast repertoire of instruments for fine-grained analysis of the cognitive structures held by individuals. However, these methods are usually based on interviews, observations, and analysis of performance protocols. Designed for laboratory studies, such techniques are time-consuming, slow, and not always appropriate for assessment purposes in realistic instructional settings, especially in adaptive e-learning environments that need rapid measures of proficiency suitable for monitoring changes in a learner's knowledge base in real time.

Multiple studies in cognitive science demonstrate that organised schema-based knowledge structures in long-term memory are the most important factor in proficient expert performance (e.g., Chi, Feltovich, & Glaser, 1981; Larkin, McDermott, Simon, & Simon, 1980). These cognitive constructs effectively reduce or eliminate severe processing limitations and fundamentally alter the characteristics of cognitive

performance. In the absence of relevant schematic knowledge, people usually use various search techniques to solve problems (e.g., the trial-and-error method or means-ends approach). These methods are time-consuming and may impose a heavy cognitive load (Sweller, 1988). In contrast, proficient expert performance is characterised by rapid retrieval and application of previously acquired schematic solution procedures from the expert's knowledge base. The availability and levels of acquisition of such knowledge structures in long-term memory are the major characteristics of a learner's proficiency in a specific domain.

If a major aim of instruction is the acquisition of organised knowledge structures (schemas) in a learner's long-term memory, then measuring the degree of acquisition of such knowledge structures should be a major purpose of diagnostic assessment. "We need to have assessment instruments that capture the complexity of understanding and allow us to make inferences about the degree to which individuals have put their knowledge together into useful cognitive structures" (Marshall, 1995, p. 448). Possible schema-based approaches to the assessment of learner knowledge structures and design of test items have been discussed in recent years (Marshall, 1993, 1995; Singley & Bennett, 2002).

The above examples demonstrating the diagnostic shortfalls of traditional testing approaches show that evidence based on these results, which are distanced from the cognitive processes involved in test performance, when used for cognitive diagnosis purposes, could be flawed because of contamination by unaccountable cognitive factors. For example, students may use problem-search strategies (consistently or partially, during some solution stages) instead of expected knowledge-based procedures. Sophisticated measurement models could potentially be developed to account for such factors. However, collecting valid evidence that is directly relevant to the assessed cognitive constructs is a better and more economical way of ensuring the accuracy of diagnostic assessment.

Cognitive Base of Rapid Assessment Approaches

Working memory is the locus of our conscious cognitive processing and represents a "window" through which we can directly access cognitive processes and observe authentic cognitive structures in action. Working memory operation involves both high-level cognitive processing (e.g., constructing and integrating mental representations) and short-term maintenance of information involved in those processes (see Miyake & Shah, 1999, for an overview). When we process unfamiliar information, our working memory is very limited in capacity and duration. According to Miller (1956), we can remember no more than approximately seven units of information without rehearsing. Working memory capacity may easily become overloaded when many elements of information are processed concurrently.

Long-term memory contains our knowledge base in the form of hierarchically organised, domain-specific structures (schemas), and is practically unlimited in both capacity and duration. These long-term memory knowledge structures are capable of reducing working memory load significantly by allowing multiple elements of

information to be treated as a single chunk. Experts in a domain, who have their knowledge organised in schematic structures in long-term memory, have no set limits on the amount of information they can process in their working memory (Sweller, 2003). Because schematic knowledge structures held in long-term memory define the key characteristics of working memory, the immediate content of working memory during complex cognitive activities could be traced in order to assess levels of learner proficiency in a domain. An assessment approach that emphasises the immediate content of working memory in realistic learning environments while students approach a situation, try to make sense out of presented information, or solve a task may provide a good indicator of their schematic knowledge base in the domain.

Experts are capable of activating and bringing into working memory their high-level, often automated, long-term memory structures, which encapsulate large amounts of information. These structures allow effortless processing of this information, thus extending working memory limits. For example, in the case of algebra equations, the working memory processing limits of a more proficient student who can immediately give the final answer to the equation $2x + 3 = 4$ are extended considerably in comparison with the processing limits of a novice student who applies a random search solution process, or even an intermediate-level student who applies a detailed step-by-step fine-grained solution procedure. Therefore, the suggested approach would in effect assess the extent to which working memory limits have been altered by schematic knowledge structures held in long-term memory.

Long-term memory structures associated with the content of working memory form a long-term working memory structure (Ericsson & Kintsch, 1995) that is capable of holding large amounts of information due to chunking effects provided by the associated long-term memory schemas. For example, when reading a text, we construct and continuously update a mental representation of the text in our working memory, retrieving associated components of our knowledge base from long-term memory. This integrated cognitive construct represents the current content of our long-term working memory. Due to associations with established long-term memory structures, it is durable enough to survive, for example, temporary interruptions in reading (see Kintsch, 1998, for details of a theory of reading comprehension).

If a learner encounters a test task in a familiar domain, her or his immediate approach to this task is based on available schematic knowledge structures. These structures are rapidly activated and a corresponding long-term working memory structure is created. This construct is sufficiently durable and interference-proof to allow ample time for a practical assessment procedure. For example, students will be able to record or otherwise register their responses in a suitable format. Thus, it is practically feasible to determine the content of a student's long-term working memory using an appropriately formulated set of cognitive tasks.

An obvious way of utilising this idea in practice is to ask students to "think aloud" as they approach a problem, inspect a diagram, or read a text (Chi,

Bassok, Lewis, Reimann, & Glaser, 1989; Ericsson & Simon, 1993; Magliano & Millis, 2003). Such an assessment procedure, however, could be time-consuming and difficult to computerise. A possible alternative method could be based on presenting each test problem to students for a limited time and asking them to indicate rapidly their first step toward the solution of the task (instead of providing a complete solution). Learners who are more proficient in the domain should be better able to immediately recognise and retrieve appropriate higher-level solution schemas than less knowledgeable learners (Kalyuga, 2003; Kalyuga & Sweller, 2004).

Because students are required to indicate only one (their first) step for each task, the amount of steps involved in the test and testing time would obviously be reduced in comparison with traditional methods. However, the most important advantage of the rapid assessment method is its ability to capture the content of a learner's working memory when she or he approaches a task. Different first-step responses would reflect different levels of acquisition of the corresponding schematic knowledge in the learner's long-term memory. More proficient problem solvers would use their high-level schemas, which would allow them to skip many intermediate steps. Novices would start their search procedure (e.g., trial-and-error method). Intermediate-level learners might be able to indicate the very first operation in the detailed step-by-step solution procedure.

Thus, the rapid assessment method may allow detection of not only the availability of necessary knowledge of schematic solution procedures, but also the levels of their acquisition. When applying a rapid assessment approach in a general case (involving both declarative and/or procedural knowledge structures), a student could be presented with an appropriate set of tasks for a limited time and required to indicate the highest level of schematic knowledge she or he is capable of retrieving immediately and applying to the presented material.

If test tasks require responses that cannot be specified exactly (e.g., verbal sequences or drawing graphical representations), recording and analysing students' rapid first-step responses in computer-based tests may not be technically plausible. An alternative rapid assessment method based on students' verifications of possible solution steps could be used in such situations. With the rapid verification approach, each task is presented to learners for a limited time. However, instead of asking students to provide their first solution step, the students are presented with a sequence of possible solution steps. Each step is presented for a limited time (e.g., a few seconds), and students are required to indicate immediately whether the step is correct or incorrect, for example by pressing a corresponding key on the computer keyboard.

The rapid verification approach allows for the design of on-line rapid tests for practically any domain and type of knowledge. In some complex and/or not well-structured domains that involve multiple-step problems, students might be able to take different routes to problem solutions. If all these problem-solving routes are identifiable, the first-step assessment procedure still could be used in both paper-based and electronic formats. However, if the number of routes is too large, the

application of a rapid assessment method may require some restrictive assumptions. In this case, it is possible to select only a limited number of steps representing different levels of schema-based problem solutions. Then the levels of proficiency could be assessed by asking participants to verify rapidly each of the sequentially presented selected solution steps.

The design procedure for a rapid assessment of learners' proficiency in any domain requires, first, the definition of the domain-specific schematic knowledge structures that guide cognitive processing. Second, a sequence of tasks that allow the collection of evidence about these knowledge structures should be constructed and a scoring procedure defined. The scoring procedure should assign specific scores for different levels of student response corresponding to different levels of proficiency in the domain. If the purpose of assessment is obtaining the overall characteristics of student performance, a classical test model could be applied to the data to generate the required statistics and rank each student's performance. To assess levels of acquisition of each schematic cognitive construct separately, an appropriate multidimensional measurement model could be fitted to the data.

The rapid schema-based approach to the assessment of levels of learner proficiency was developed and pilot-tested in several domains: simple linear algebra equations, coordinate geometry tasks, and arithmetic word problems (Kalyuga, 2004, *in press*; Kalyuga & Sweller, 2005). These areas involve relatively well-defined tasks and predictable sequences of solution steps, and the application of rapid assessment methods is straightforward. In contrast, reading tasks represent a less-structured domain where students might take different routes to achieving comprehension. The remaining sections of the paper describe an attempt to apply the rapid approach to the assessment of reading proficiency.

Cognitive Processes in Reading Comprehension

The first step in the design of a rapid assessment of reading proficiency is defining the organised knowledge structures that guide cognitive processing during reading. A reader comprehends a sentence or a textual passage by building a mental representation of the text. For example, the construction-integration theory of text comprehension (Kintsch, 1998) suggests formation of the initial propositional network, which is then integrated into a more coherent representation resulting in the formation of the overall model. The ability rapidly to construct rich mental representations of the text by effortlessly integrating newly read modules into a coherent overall structure is an important characteristic of a skilled reader.

Text comprehension as a high-level cognitive process is guided explicitly or implicitly by available long-term memory knowledge structures. These organised knowledge structures may include specific knowledge related to situations described in the text, knowledge of syntactic structures, and lower-level knowledge structures associated with the meaning of separate words, symbols, etc. During reading, available long-term memory knowledge structures (e.g., syntactical or grammatical rules) constantly interact with information in working memory.

Once an appropriate set of propositions is constructed in working memory, it is integrated into a knowledge structure stored in long-term memory. This structure then provides necessary guidance during further text processing. Because of the need to establish causal links between propositions in working memory, the process of integration of the propositions could be cognitively very demanding (Perfetti, 1985; Waters & Caplan, 2001). Due to working memory limitations, the reader can only hold several propositions simultaneously in working memory. They should be integrated into a more general representation before being misplaced as new propositions are formed. For long and structurally complex sentences, such integration may require a great deal of processing effort, especially for unskilled readers.

Syntactic rules and patterns in long-term memory, activated when a person starts reading a sentence, may substantially reduce the processing load. Due to syntactic organisation, information can be held in working memory until the rest of the sentence is processed and while other processes are performed (Carpenter & Just, 1989; King & Just, 1991). When constructing a sentence or text representation, syntactic structures allow readers to determine propositions that are important for the overall representation and avoid a cognitively inefficient random search for possible causal relations. Syntactic knowledge guides the reader in dividing a text into meaningful units and performs the function of mental processing instruction (Kintsch, 1995). It is especially important for sentences that include relationships between nonadjacent components, for example long-distance syntactic dependencies (e.g., “This is the movie that the studio believed the audience enjoyed most”). Knowledge of common syntactic structures also assists readers in directly accessing previously built representations (McElree, Foraker, & Dyer, 2003).

Cognitive processing load during sentence comprehension depends on how long the syntactic predictions and requirements must be remembered before they are fulfilled, and on the distance between an incoming word and the words to which it attaches (Gibson, 1998). These factors explain the high processing complexity of relative clauses, embedded dependencies, and multiple-nested structures (e.g., “This was the course that the student, whom the school that was criticised by the newspaper expelled, failed”). For multiple-embedded sentences, working memory load can become excessive, even for very skilled readers. However, less skilled readers may find less structurally complex sentences difficult to comprehend.

Reading skills could be assessed by the quality of mental representations that readers are capable of constructing through retrieving and applying their linguistic knowledge structures. When sentences with different levels of structural and grammatical complexity are presented for a limited time, less skilled readers might be able to make sense of only simple sentences due to the excessive working memory load associated with processing complex relations. More experienced readers would understand increasingly more complex sentences according to the level of knowledge of grammatical structures they possess. Thus, relative levels of reading expertise could be operationally defined by the level of structural complexity of the corresponding test sentences. A procedure for constructing an appropriate set of sentences will be described in the following section.

A Sequence of Tasks for Rapid Assessment of Reading Proficiency

To assess levels of reading expertise, it is possible to modify the sentence-verification approach and produce a rapid verification approach (e.g., Royer, Hastings, & Hook, 1979; Schwartz, 1984). After a participant finishes reading a sentence, she or he could be presented with a set of statements and asked to decide whether these statements are correct or incorrect. This procedure would indicate whether (and how accurately) the reader was able to derive meaning from a statement held in working memory and compare it with the previously constructed representation of the entire original sentence.

A task pattern for the assessment of reading expertise could be constructed as a sequence of sentences with gradually increasing structural complexity. The set may begin with simple sentences that are followed by compound sentences. Then, more complex sentences should follow. These sentences could contain an increasing number of dependent and/or embedded clauses. For each sentence in the set, several (e.g., four) simple statements (both correct and incorrect) should be constructed.

For example, for the sentence “The artist, who performed for the crowd that gathered to enjoy the show, left,” the verification statements could be (the first and third statements are correct): “The artist left”; “The artist enjoyed the show”; “The crowd gathered for the show”; and “The crowd left.”

The ability of a student to identify correct and incorrect statements immediately after reading increasingly more structurally complex sentences could be considered an indicator of her or his reading expertise. In order to control for the influence of domain-specific knowledge, the sentences should be based on a restricted range of common everyday experiences that are familiar to all participants and do not require any additional specialised knowledge.

Processing complex multiple-nested syntactic structures might easily exceed limited working memory capacity which, according to some estimations, allows no more than four local thematic violations (Gibson, 1998) or two sentence nodes or clauses to be attended to at one time (Kimball, 1973). Level of reading proficiency could be associated with the level of cognitive difficulty of the sentences that the reader is able to comprehend effortlessly and, consequently, for which she or he can rapidly verify the corresponding statements. The level of cognitive difficulty (D) of each sentence is formally defined as the larger of two numbers: (1) the number of local thematic violations; and (2) the number of sentence nodes or clauses to be considered simultaneously. That is:

$$D = \max (\text{local thematic violations}; \text{sentence nodes or clauses to consider at one time})$$

Sentences with D values of 1, 2, or 3 are associated with corresponding levels (1, 2, or 3) of reading proficiency. Sentences with D values of 4 or higher are associated with reading proficiency level 4.

For example, the simple sentence “The noise of the crowded city was replaced by the silence of the countryside” is an example of the first level sentence ($D = \max [0, 1] = 1$). A complex sentence with an embedded clause such as “The student whom the class liked finished the project” represents the second level of cognitive difficulty

($D = [2, 2] = 2$), because it has two local thematic violations (“the student” and “whom”) and two sentence nodes to parse (“the student finished” and “the class liked”). Complex sentences with two embedded clauses such as “The student, whom the teacher recommended, reported to the trainer who organised the school team” are level 3 sentences ($D = [2, 3] = 3$), with two local thematic violations (“the student” and “whom”) and three sentence nodes to parse (“the student reported,” “the teacher recommended,” and “the trainer organised”). Complex sentences with double-embedded clauses such as “The fact that the student whom the teacher praised had failed the course worried the principal” represent the fifth level of difficulty ($D = [5, 3] = 5$). It has five local thematic violations (“the fact,” “that,” “the student,” “whom,” and “the teacher”) and three sentence nodes to parse (“the fact worried,” “the student failed,” and “the teacher praised”).

The most difficult sentences are sentences with multiple-embedded clauses, for example, “This was the course that the student, whom the school that was criticised by the newspaper expelled, failed” (see Figure 1). This sentence has the difficulty level 7 ($D = [7, 4] = 7$), with seven local thematic violations (“the course,” “that,” “the student,” “whom,” “the school,” “that,” and “the newspaper”) and four sentence nodes to parse at one time (“the course was,” “the student failed,” “the school expelled,” and “the newspaper criticised”). This sentence could be extremely difficult to comprehend within a limited amount of time, unless the reader is highly proficient.

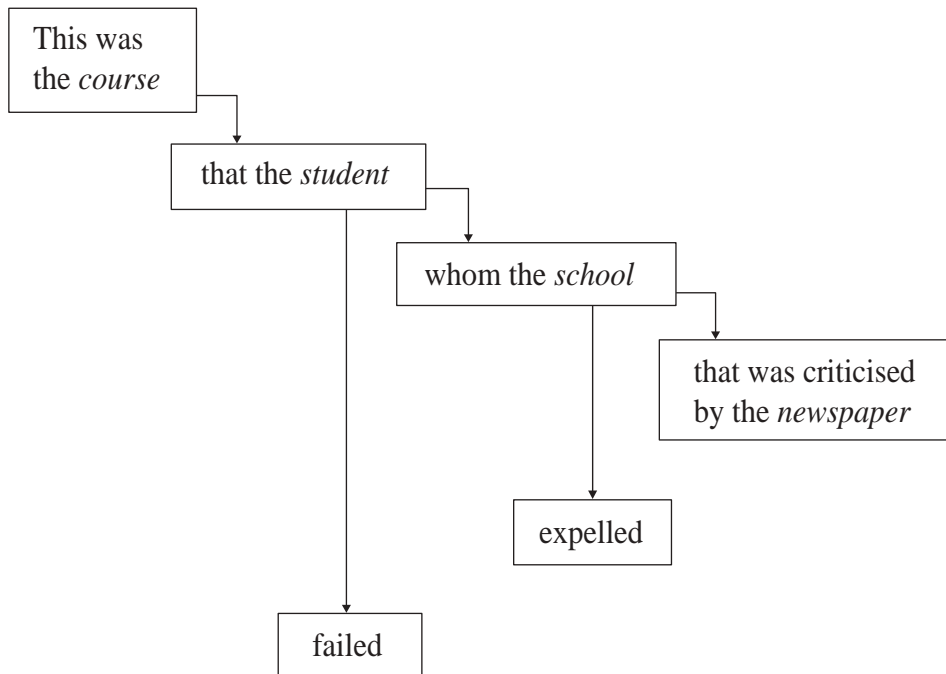


Figure 1. An example of the complex multiple-nested syntactic structure

Method

The evidence regarding reading proficiency that can be obtained from the described pattern of tasks is based on the readers' abilities to process cognitively difficult sentences under heavy working memory load. Students in this study were presented with a sequence of sentences, one at a time for a limited time, arranged according to level of difficulty. After reading each sentence, each student needed to verify immediately a set of statements related to the sentence.

Participants

A sample of 34 Year 7 students (average age of around 13) from a Sydney boys' school participated in the experiment. The students had not been exposed to the sentences used in the study prior to the experiment.

Materials and Procedure

The experiment was conducted in a realistic class environment. To find out if the rapid test would demonstrate sufficient predictive validity by correlating highly with traditional measures of reading skills, a traditional reading comprehension test was also administered to the same students. In the traditional test, students were required to provide answers to multiple-choice questions about textual passages they read. The traditional paper-based test was constructed from items used in Australian school English competitions in previous years, and included eight text passages, each followed by a set of four to six multiple-choice questions (total of 42 questions). Self-paced tasks were used in both tests to determine actual time reductions associated with rapid testing. All participants were tested simultaneously in one session.

The rapid test was computer-based with automatic recording of student responses and response times. A sequence of 18 sentences with gradually increasing levels of cognitive difficulty was constructed for the test. Each sentence was displayed for a limited time that was set using the approximate rate of one second per word. This rate had been established in pre-trials and allowed ample time for students to read each sentence once. After a sentence disappeared from the screen, four brief statements (correct and incorrect) about this sentence were displayed on the screen, one statement at a time. The pre-test instruction to students included the following statement:

On each of the following pages, you will see a sentence. You will be allowed a limited time to read it.

Following each sentence, several related brief statements will be presented. For each statement, you have to press immediately a green key (on your right) if the statement is correct or a red key (on your left) if the statement is incorrect.

Appropriate keys on computer keyboards were marked with green and red stickers. Each response was automatically scored 1 or 0 depending on whether the verification was successful or not. Before the test, an exercise sample based on a sentence

not used during the test was presented to students to “coach” them in responding quickly. During the exercise, students could see how rapidly they were expected to respond. If a student did not respond within a set time interval of several seconds, she or he was asked to respond faster next time until, eventually, the responses became rapid enough.

Results and Discussion

The variables under analysis were: traditional test time (time in seconds each learner spent on reading all eight passages and answering all 42 questions); rapid test time (time each learner spent on reading sentences and verifying all 72 statements); test scores (number correct) for the traditional test; and test scores (number correct) for the rapid test. Means and standard deviations are displayed in Table 1.

There was a significant difference between test times for the traditional and rapid tests ($t[32] = 25.8, p < .0001$), with a reduction factor of 3.8 (from 28.1 minutes to 7.4 minutes). The Pearson product-moment correlation between scores for the traditional and rapid tests was .66 ($p < .01$), suggesting a sufficiently high degree of predictive utility for the rapid test. Estimates of reliability coefficients (as indicators of the internal consistency of the tests) using the Kuder-Richardson formula (KR20) were .27 for the traditional test and .73 for the rapid test. Low reliability of the traditional test might be a consequence of a heterogeneous aggregation of different factors. The single-factor linear model (within which the KR20 is applicable) may not be suitable for data that have a multidimensional structure, and a low internal consistency could be expected. Nevertheless, the traditional test reflects a combination of relevant factors and is still a valid external measure for estimating the predictive utility of the rapid test. All rapid test items reflected the level of integration of propositions and thus were more homogeneous and less sensitive to the underlying multidimensionality. (In a separate trial of a similar rapid reading test with 71 secondary school students, an estimate of its reliability was .85.)

To evaluate the diagnostic potential of the rapid test, we compared the total scores for three sequential groups of sentences (six sentences in each group) reflecting

Table 1. Mean test time and scores by testing conditions

Variable	Testing conditions	
	Traditional test	Rapid test
Test time (seconds)		
<i>M</i>	1,684	441
<i>SD</i>	256	51
Total test score (% correct)		
<i>M</i>	29.8 (71%)	40.3 (56%)
<i>SD</i>	2.7	5.5

different levels of sentence complexity. As expected, there was a gradual decrease in scores with increase in sentence complexity (means and standard deviations, correspondingly, 16.5/2.2, 14.0/2.9, and 11.8/3.3) with a significant overall difference ($F[2,64] = 3.18$, $MSe = 6.83$, $p < .05$). The test allowed us to identify those students who had difficulties with structurally more complex sentences (e.g., score $< 50\%$ correct for the corresponding group of sentences) but were able to handle relatively more simple sentences (score $> 50\%$ correct). The data also indicated the level of sentence complexity at which difficulties began to appear.

Such diagnostic data can be used to plan meaningful instructional interventions for less proficient readers. For instance, various complex sentence structures could be explicitly demonstrated to learners using worked examples with hierarchical diagrammatic sentence representations (e.g., similar to that presented in Figure 1, complemented by colour-coding of different levels of hierarchy). Each example could be followed by a set of exercises in analysing the structure of analogous complex sentences and making sense out of them.

Means and standard deviations for student response times were $M = 2.83$ and $SD = .75$ for correct statements, and $M = 3.22$ and $SD = .73$ for incorrect statements. There was a significant difference between response times for correct and incorrect statements ($t[32] = 2.94$, $p < .01$). This difference was not unexpected, considering that the response to a correct statement might be based on the recognition of a relevant part of the previously constructed propositional structure of the sentence. The verification of an incorrect statement might require more extended search for a possible segment in the propositional structure of the sentence. Individual performance records showed that there were several participants whose average response times for incorrect statements were lower or comparable to their average response times for correct statements. Remarkably, for these individuals the records also showed unexpected increases (or absence of decreases) in score values with the increase in complexity of sentences. These irregularities could well be a consequence of the students guessing their responses (or responding randomly). Thus, comparing average response times for correct and incorrect statements could potentially be used to identify students who guess their responses during computer-based rapid reading tests. Appropriate instructions for these students could be displayed on the screen.

Conclusion

To be cognitively efficient, instructional techniques need to change with alterations in learner proficiency in a domain. To match instructions to individual learners, it is necessary to have appropriate means of measuring levels of their proficiency. The required measures should be rapid enough to monitor changes in the learner's knowledge base during instruction (e.g., during a single instructional session), especially in e-learning adaptive environments. On the other hand, the measures should be sensitive to different levels of acquisition of organised knowledge structures. Most traditional testing methods are not suitable for this purpose.

Organised schematic knowledge structures in long-term memory effectively define the characteristics of our working memory. Therefore, monitoring the content of working memory during complex cognitive activities can provide a means of diagnosing levels of learners' knowledge-based proficiency. This paper describes an attempt to apply the rapid assessment approach to measuring levels of reading proficiency. Contrary to some other areas (e.g., mathematics) this domain does not allow precise specification of the distinct schematic knowledge structures used in reading comprehension. Instead, a rather coarse-grained operational description of proficiency is suggested. Different levels of reading proficiency could be defined according to the level of cognitive difficulty of the test sentences that the reader is able to comprehend effortlessly in a limited period of time.

Using the suggested rapid assessment tool for practical purposes in realistic educational settings would require the construction and calibration of sets of items suitable for different learner populations (e.g., different age groups). When appropriately constructed and trialled, these instruments could be embedded into computer-based instructional packages as monitoring diagnostic tools for tailoring instructional procedures and materials to changing levels of student proficiency in cognitively optimal ways and for developing adaptive e-learning systems.

However, before practical implementations of cognitively optimised adaptive technologies of instruction can be developed, more research studies are required. The generality of the rapid assessment methods should be established by applying them in other domains. The reliability of such tests and their correlations with traditional performance measures in these domains should be estimated as indicators of the predictive validity of the approach. Most applications of rapid diagnostic techniques have been concerned primarily with the assessment of procedural knowledge. The same approach could be applied in diagnostic assessment of declarative schematic knowledge structures.

The usability of the rapid schema-based assessment method as a means of real-time tailoring of instruction to changing levels of learner proficiency was experimentally tested only in simple computer-based algebra tutorials (Kalyuga & Sweller, 2005). In those studies, the rapid test was used to select the appropriate initial level of detail in instructional explanations, as well as for monitoring learners' progress during instruction and real-time adjusting of the level of instructional guidance. The studies demonstrated that learning was enhanced by adapting instruction to learner levels of proficiency, and they provide strong evidence for the importance and usability of the rapid testing techniques. These initial results need to be generalised to other knowledge domains. For example, the described method of rapid assessment of reading proficiency could be experimentally tested as a diagnostic tool in computer-based language tutorials to tailor instructional materials to changing levels of language proficiency. The technique would allow the incorporation of cognitive load as an essential factor in the learning task selection procedure, and could be considered as one means of enhancing the adaptability of computer-based and e-learning instructional systems.

References

- Carpenter, P. A., & Just, M. A. (1989). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 31–68). Hillsdale, NJ: Erlbaum.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanation: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.
- Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.
- Kalyuga, S. (2003). Rapid assessment of learners' knowledge in adaptive learning environments. In U Hoppe, F. Verdejo, & J. Kay (Eds.), *Artificial intelligence in education: Shaping the future of learning through intelligent technologies* (pp. 167–174). Amsterdam: IOS Press.
- Kalyuga, S. (2004, April). *Rapid dynamic assessment of expertise to manage cognitive load during instruction*. Paper presented at the 2004 Annual Meeting of the American Educational Research Association (AERA), San Diego, CA.
- Kalyuga, S. (in press). Instructing and testing for expertise: A cognitive load perspective. In A. V. Mitel (Ed.), *Focus on educational psychology*. Hauppauge, NY: Nova Science Publishers.
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). Expertise reversal effect. *Educational Psychologist*, *38*, 23–31.
- Kalyuga, S., Chandler, P., & Sweller, J. (2001). Learner experience and efficiency of instructional guidance. *Educational Psychology*, *21*, 5–23.
- Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology*, *96*, 558–568.
- Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology, Research and Development*, *53*, 83–93.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*, 15–47.
- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580–602.
- Kintsch, W. (1995). How readers construct situation models for stories: The role of syntactic cues and causal inferences. In M. A. Gernsbacher & T. Givon (Eds.), *Coherence in spontaneous text* (pp. 139–160). Amsterdam: John Benjamins Publishing.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Larkin, J., McDermott, J., Simon, D., & Simon, H. (1980). Models of competence in solving physics problems. *Cognitive Science*, *4*, 317–348.
- Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction*, *13*, 251–283.
- Marshall, S. P. (1993). Assessing schema knowledge. In N. Frederiksen, R. J. Mitlevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests* (pp. 155–179). Mahwah, NJ: Lawrence Erlbaum Associates.
- Marshall, S. (1995). Some suggestions for alternative assessments. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively diagnostic assessment* (pp. 431–453). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*, 67–91.

- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.
- Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge, UK: Cambridge University Press.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Pellegrino, J., Chudowsky, N., & Glaser, R., (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Royer, J. M., Hastings, C. N., & Hook, C. (1979). A sentence verification technique for measuring reading comprehension. *Journal of Reading Behavior*, 11, 355–363.
- Schwartz, S. (1984). *Measuring reading competence: A theoretical-prescriptive approach*. New York: Plenum Press.
- Singley, M. K., & Bennett, R. E. (2002). Item generation and beyond: Applications of schema theory to mathematics assessment. In S. H. Irvine & P. C. Kyllonen (Eds.), *Item generation for test development* (pp. 361–384). Mahwah, NJ: Lawrence Erlbaum Associates.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257–285.
- Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–266). San Diego, CA: Academic Press.
- Waters, G., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging*, 16, 128–144.

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