



TOEFL[®]

Monograph Series

MS - 4
SEPTEMBER 1996

*Assessing Second Language
Academic Reading from a
Communicative Competence
Perspective: Relevance for
TOEFL 2000*

Thom Hudson



**Assessing Second Language Academic Reading from a
Communicative Competence Perspective: Relevance for TOEFL 2000**

Thom Hudson

**Educational Testing Service
Princeton, New Jersey
RM-96-6**



Educational Testing Service is an Equal Opportunity/Affirmative Action Employer.

Copyright © 1996 by Educational Testing Service. All rights reserved.

No part of this report may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher. Violators will be prosecuted in accordance with both U.S. and international copyright laws.

EDUCATIONAL TESTING SERVICE, ETS, the ETS logo, TOEFL, the TOEFL logo, and TSE are registered trademarks of Educational Testing Service.

Foreword

The TOEFL® Monograph Series features commissioned papers and reports for TOEFL 2000 and other Test of English as a Foreign Language program development efforts. As part of the foundation for the TOEFL 2000 project, a number of papers and reports were commissioned from experts within the fields of measurement and language teaching and testing. The resulting critical reviews and expert opinions were invited to inform TOEFL program development efforts with respect to test construct, test user needs, and test delivery. Opinions expressed in these papers are those of the authors and do not necessarily reflect the views or intentions of the TOEFL program.

These monographs are also of general scholarly interest, and the TOEFL program is pleased to make them available to colleagues in the fields of language teaching and testing and international student admissions in higher education.

The TOEFL 2000 project is a broad effort under which language testing at ETS will evolve into the 21st century. As a first step in the evolution of TOEFL language testing, the TOEFL program recently revised the Test of Spoken English (TSE®) test and announced plans to introduce a TOEFL computer-based test (TOEFL CBT) in 1998. The revised TSE, introduced in July 1995, is based on an underlying construct of communicative language ability and represents a process approach to test validation. The TOEFL CBT will take advantage of the new forms of assessments and improved services made possible by computer-based testing while also moving the program toward its longer-range goals, which include

- the development of a conceptual framework that takes into account models of communicative competence
- a research agenda that informs and supports this emerging framework
- a better understanding of the kinds of information test users need and want from the TOEFL test
- a better understanding of the technological capabilities for delivery of TOEFL tests into the next century

It is expected that the TOEFL 2000 efforts will continue to produce a set of improved language tests that recognize the dynamic, evolutionary nature of assessment practices and that promote responsiveness to test user needs. As future papers and projects are completed, monographs will continue to be released to the public in this new TOEFL research publication series.

TOEFL Program Office
Educational Testing Service

Abstract

This paper examines issues involved in the assessment of academic reading from a communicative proficiency perspective, particularly these issues as they are involved in the context of the Test of English as a Foreign Language TOEFL 2000 project. The first part of the paper presents the areas that concern the assessment of academic reading ability by briefly examining notions relevant to communicative competence and how these notions might relate to academic reading. This discusses the ways in which models of communicative competence have broadened the view of what knowledges are necessary in order to use a language and language skill areas such as academic reading. The paper notes that the frameworks that have evolved have included such areas as grammatical competence, organizational competence, illocutionary competence, and pragmatic competence. Such views of competence and performance are important in language assessment in that in addition to broadening views of language and language ability, they offer some means for explaining the extent to which a person might vary in language performance across tasks or contexts. The stress in communicative competence perspectives on language use reflects an emphasis on the importance of viewing language in context.

The paper employs a broad view of academic reading that includes key aspects of (a) automaticity in word and sentence processing, (b) content and formal background knowledge, (c) strategic and metacognitive skill application, and (d) reading purpose and context. By acknowledging the important role played by each of these aspects, the paper proposes that future assessment take the ecology of academic reading into account. In short, reading simultaneously involves both psychological and sociological processes.

The paper concludes with implications for academic reading assessment, paying particular attention to the four validity components of (a) construct validity, (b) value implications, (c) relevance/utility, and (d) social consequences. The implications involve the potential needs to (a) expand beyond the multiple-choice or other selected-response formats, (b) incorporate some multiple-choice or other selected-response formats in order to balance general and context-specific tasks, (c) view reading assessment as, at least in part, task based, (d) involve thematically based texts and develop adaptive types of tests, either via computer or through selection of texts and tasks by the examinee, (e) view score reporting and interpretation as reflecting real-world academic reading tasks, and (f) integrate reading assessment with other language skills based on the "literacy" task that is being encountered.

Table of Contents

	Page
Introduction	1
Communicative Competence and Reading	2
Review of the Literature on the Construct of Academic Reading	3
Models of the L1 and L2 Reading Processes	4
Implications for the Assessment of Academic Reading	9
Conclusion	13
References	14

Introduction

This paper focuses on issues involved in the assessment of academic reading from a communicative proficiency perspective in the context of large-scale assessment for academic selection. Specifically, it addresses these issues for the context of the TOEFL 2000 project. The first part of this paper will present the areas that are of concern in the assessment of academic reading ability. This will involve a brief examination of ideas relevant to communicative competence and a discussion of how these issues might relate to academic reading. Issues relevant to defining the construct of academic reading will then be discussed, concentrating on aspects involved in communicatively assessing academic reading. Then, the implications of reading assessment discussion will be explored, and the possibilities for integrating the various language skills will be discussed.

The approach taken here involves broadening the view of academic reading to include key aspects of (a) automaticity in word and sentence processing, (b) content and formal background knowledge, (c) strategic and metacognitive skill application, and (d) reading purpose context. By acknowledging the important role played by each of these aspects, the paper proposes that future assessment take the ecology of academic reading into account. In short, reading simultaneously involves both psychological and sociological processes. In order for assessment to more fully reflect the construct of academic reading, it will need to address the full range of applied reading processes.

Communicative Competence and Reading

Over the past several years, the language analysis and assessment community has increasingly emphasized the role of *communicative competence* and the appropriate means for its assessment. This emphasis has developed from theories of language that view language *use* as a primary component that must be addressed in studies of language analysis, acquisition, and education. Communicative competence was proposed by Hymes (1972) and Campbell and Wales (1970) as an alternative to the strong view posited by Chomsky that linguistic competence was limited to knowledge of grammatical rules. Communicative competence perspectives, in contrast, view the role played by context of discourse beyond sentential constructions as essential to any understanding of competence and performance. Thus, models of communicative competence have broadened the view of what knowledges are necessary in order to use a language (Bachman & Savignon, 1986). Current models include knowledge of language functions and knowledge of language contexts, as well as knowledge of grammar. The frameworks that have evolved have included grammatical competence, sociolinguistic competence, and strategic competence (Canale & Swain, 1980), or grammatical competence and textual competence, organized under a rubric of organizational competence with illocutionary competence and sociolinguistic competence under pragmatic competence (Bachman, 1990). The Bachman view of competence and performance is important in language assessment in that it offers some mechanism for explaining the extent to which an individual's language performance might vary across tasks. In addition to the trait factors of linguistic competence, pragmatic competence, and strategic competence, the Bachman model adds skill factors (such as psychophysiological mechanisms, mode, and channel) as well as method factors (such as the language-use situation, the amount of context, the distribution of information, the type of information, and the response mode). Skehen (1991) points out that including these factors may help address the competence versus performance issue. Basic competences are concerned with general language abilities, while skill and method relate the general abilities to real-world language performance.

The stress in communicative competence perspectives on language use reflects an emphasis on the importance of viewing language in context. As Savignon (1991) has noted, communicative language teaching embraces both the goals and the processes of learning, thus providing a focus on social interaction and language acquisition. This view is consistent with Green's (1987) observations about how reading has traditionally been addressed. She points out that most prior views of reading have either seen reading as an *object of* instruction or as a *medium for* instruction in other curricular areas. Such a dichotomization has led to a focus in reading instruction and assessment on the skills, strategies, and processes individual readers acquire and has tended to decontextualize the activities of reading and literacy from the social and instructional context in which they are embedded. This perspective of reading and comprehension as active and interactive is consistent with the notion that comprehension models need to incorporate both rules and strategies (van Dijk, 1985). Strategies are applied to rules in order to reach an adequate text interpretation, and strategies are defined, in part, in terms of the reader's purpose and determined need. Thus, the end goal of the reading process will determine what strategies are applied to which rules in order to determine such reflections of ability as depth of processing or time spent in reading. Current research has come to view ability as variable and highly dependent upon context and purpose (Savignon, 1991).

These are important issues for language assessment. As Canale and Swain (1980) noted, communicative testing “. . . must be devoted not only to what the learner knows about the second language and about how to use it [competence] but also to what extent the learner is able to actually demonstrate this knowledge in a meaningful communicative situation” (p. 32). Thus, in expanding the views of reading to incorporate such concepts as process, goal, context, and variability of ability, it may be that we expand the application of reading assessment to literacy in academic settings, rather than restricting reading assessment to traditional concerns with automaticity of reading skills. This view of reading as academic literacy may provide valuable information for selection decisions at academic institutions, where what a potential candidate can do with the language is as important as that candidate’s underlying competence. Likewise, from a communicative competence framework, it is important to recognize that a candidate should be allowed to demonstrate the ability to apply reading skills to a task in purposeful sociocultural context.

Review of the Literature on the Construct of Academic Reading

Defining proficient reading in the academic context is a difficult endeavor. One of the problems we have with reading research and model building is that since the component processes are unobservable, they are postulated constructs. These processes are inferred from the results of behaviors. An additional problem with building reading models is that the models that emerge are always products of history, limited by the theories and instruments available. Given this historical background in which models operate, frameworks change relatively rapidly, so it is important to maintain a healthy atheoretic scepticism. Over several decades, however, a general consensus (though certainly not unanimity) about several issues surrounding reading has emerged.

There is general consensus that reading involves the interaction of a vast array of processes, knowledges, and abilities. These include basic decoding processes such as grapheme recognition, lexical access, phonological representation, and linguistic structure processing, as well as higher order cognitive processes such as the application of background knowledge, processing strategies, text structure understanding, and some aspects of vocabulary knowledge. Reading also involves interactional processes such as the application of evaluative skills, use of metacognitive knowledge, and self-monitoring. However, a major issue in reading literature relates to the specific relationships between these components. As will be explained in more detail, some researchers have advocated a bottom-up approach, others a top-down approach, and others an interactive approach to explanations of reading success and failure. Early theories viewed reading as a bottom-up process in which the reader constructs meaning in a sequential manner from letter, words, and sentences (Gough, 1972; LaBerge & Samuels, 1974). Other theories of reading stressed that the efficient reader makes the fewest text processings, since that reader predicts the meaning of the text by applying knowledge of the world and language (Goodman, 1967; Smith, 1971). More recent views acknowledge a strong interplay between both of these processes.

Current theories emphasize the interactive nature of reading and accept the fact that good readers make extensive use of printed information. In the interactive approach, some have argued in favor of an interactive system strongly constrained by bottom-up processes, and have consequently emphasized linguistic processing (Grabe, 1991; Perfetti, 1992; Rayner & Pollatsek, 1989; Stanovich, 1990). Others

have tended to place less emphasis on the bottom-up linguistic components and more emphasis on the role of top-down processing (Henk & Helfeldt, 1987; Hill & Parry, 1992; Hudson, 1991; Johnston, 1984; Rumelhart, 1977; Valencia & Stallman, 1989). Which emphasis is given seems often to be influenced by how the researchers address — explicitly or implicitly — the four issues of (a) how uniform the reading process is within any particular individual reading across contexts; (b) the importance of background and culture in reading and learning to read; (c) the extent to which reading skills are implicationally ordered in their acquisition and application as opposed to being broadly overlapping and compensatory in nature; and (d) the extent to which an individual may avoid close linguistic processing and still comprehend a message. These issues will continually arise in the discussion that follows. Essentially, the differences in emphasis relate to whether the particular researchers adopt a primarily psycholinguistic or a primarily sociolinguistic orientation. Additionally, as Grabe (1991) points out, not all discussions of *interactive* approaches use the term in the same way. Some use the term to represent an interaction between the reader and the text, while others use the term to represent an interaction between the different skills, that is, as the interaction of the entire web of cognitive assets the reader has at his or her disposal. This paper, in taking a view that context is extremely important, uses the term *interaction* to include both of the above uses.

In general, then, success or failure in reading performance can be addressed in terms of the interactions between the reader's (a) *automaticity*, the extent to which the performance of procedures no longer requires large amounts of attention; (b) *content and formal schemata*, the reader's mental representations of facts and skills; (c) *strategies and metacognitive skills*, the reader's strategies for monitoring the selection and application of actions; (d) *purpose*, the goal striven for by the reader; and (e) *context*, the interactional environment in which the reading activity takes place. Thus, regardless of the approach to reading performance, there is a need to indicate how the interactive processes in reading involve both the underlying cognitive processing and the purpose or contextual aspects of reading. In the first process, the reader interacts directly with the text to gain meaning; in the second, the reader produces some artifact. These are not separate entities in that they are both operating at the same time. They simply reflect different perspectives employed to explain the activity of what an individual does in the reading process. One perspective focuses on "skill" application while the other focuses on the *product* of the application of the skill. It is the interaction between these two perspectives that provides part of the basis for a change in assessment procedures. A recognition that reading assessment involves a balance between the two perspectives has led, in part, to the interest in communicatively oriented assessment.

Models of the L1 and L2 Reading Processes

As noted above, reading models have changed over the years. Currently, however, researchers generally agree that reading is interactive and compensatory (Stanovich, 1980). As with language in general, reading is not the sum of skills, that are identified with reading, such as skimming, scanning, vocabulary identification, and reading for main idea. What seems reasonably clear from the research is that skills are difficult to define in practice (Snow & Lohman, 1993; Alderson, 1990), and when they are identified they are not ordered implicationally (Hudson, 1993). Rather, skills appear to cover wide bands of overlapping abilities. These abilities range from local text recognition and processing to broader text interpretation and use strategies. The following discussion will present the abilities in first- and second-language reading that might be of importance in assessing academic language proficiency.

Automaticity in Word and Sentence Recognition. Current research emphasizes reading comprehension's dependence on efficient lexical access (Pollatsek & Rayner, 1990). Word recognition and decoding abilities have been shown to have strong relationships to first-language reading comprehension (Cunningham, Stanovich, & Wilson, 1990). Much of this development in the various reading models has come from research that has shown that words, sentences, and other strings of words that are processed in isolation are processed more slowly and remembered with less accuracy than those processed in context. The research has shown that, in general, context of reading and processing plays an important role in reading comprehension and memory. Balota (1990), for example, has reported that readers recognize words and strings that are primed by related words faster than they recognize words primed by unrelated words. For example, when subjects were presented with the pairs of words "dog – cat" versus "pen – cat," "cat" was recognized faster in the first instance than in the second. Additionally, some of the features of fluent reading are manifested in the time readers spend looking at, or fixating on, particular words (Beck & Carpenter, 1986; Just & Carpenter, 1987). Much of this research observed the frequency and duration of eye fixations during reading. This research showed that a large proportion of the words in a text are attended to by the reader, a much larger proportion than would be predicted by strong top-down advocates (Just & Carpenter, 1987; Pollatsek & Rayner, 1990). It is clear, then, that automatic and efficient word recognition is required for most skilled reading (Beck & Carpenter, 1986). In academic contexts where large amounts of text must be processed, this efficiency in reading must be considered to be of extreme importance. Likewise, efficient syntactic processing and vocabulary are important for efficient academic reading. Rayner (1990) points out the important interaction between syntax and lexicon in reading, and Cunningham, Stanovich, and Wilson (1990) demonstrate high correlations between reading and vocabulary. Haynes and Carr (1990) show a strong relationship between grammatical proficiency and reading speed. They claim that this indicates a large syntactic processing component in reading. Thus, the lower-level linguistic skills are essentially enabling abilities for much of academic reading.

Given the important role of efficient lexical and syntactic processing in reading comprehension, it is important for methods of assessing academic reading to represent lexical and syntactic levels appropriate to academic texts. Further, readers who lack sufficient vocabulary and/or efficient automatic text processing abilities are likely to have poorer chances for academic success than those with a high degree of ability.

Context and Formal Schema. The role of background knowledge in successful academic reading involves the application of content schemata related to the topic of the passage as well as the application of formal knowledge of how texts are organized.

Content Schema. The effect of prior knowledge on comprehension involves the theoretical notion of *schema*. As noted by Marshall (1993), schema theories have tended to focus either on schemata as abstract mental structures that serve as frameworks with slots that may be filled by incoming text information (Beck & McKeown, 1989), or as the actual knowledge of a concept that is contained in a schema (Gick & Holyoak, 1983). However, Marshall notes the need to incorporate both an architectural point of view of schema as well as a content point of view. Sharkey (1990) approaches schemata as mechanisms for collecting and representing norms, and thus presents an unmarked script that is helpful but not sufficient for comprehension. Specifically, he notes that scripts may be underspecified. It is clear

that, regardless of how the term is used, reading is not accurately represented in the traditional encoding/decoding model of communication represented below (Sperber & Wilson, 1986):

message--> encoder--> channel--> decoder --> message

This view assumed that the reader successfully decoded a message that originated from some encoder with a precise message. This is a strong view entrenched in Western thought. However, the view presents meaning as something contained within a text rather than the result of the reader determining how to make a representation of meaning (Zamel, 1992). Not every person can decode every properly constructed linguistic message, nor will every person derive the same mental representation from decoding such a properly constructed message. For example, if someone had no knowledge or framework for cooking, that person could not interpret the following sentence:

“Beat the eggs until firm.”

It is the reader’s ability to apply the appropriate meanings of “beat” and “firm” that allows comprehension. Readers with differing familiarity with cooking will likely have differing representations of what “beat” means and what constitutes “firm.”

The interaction of schemata with text is essential for understanding reading comprehension and the reading process. In their discussion of first-language reading, Anderson and Pearson (1988) note that (a) “comprehension” means the interaction of new information with old knowledge; (b) information from a new story that fits with the reader’s prior knowledge is recalled while other details are omitted or rationalized with meaning not in the original; and (c) the author of a story suggests information to the reader and, if it is a new or unfamiliar concept, comprehension resembles a general problem-solving activity. Further, Symons and Pressley (1993) have shown that prior knowledge affects the way in which a reader searches through a text for information. Anderson and Pearson (1988) claim that an adequate account of the structure of the role of prior knowledge, or schemata, will include: (a) information about the relationships among the components; (b) a major role for inference; (c) acceptance that during language comprehension, people probably rely on knowledge of particular cases as well as abstract and general schemata. Moreover, work in cultural background, content knowledge, language skills, and reading comprehension has shown that schemata are important for reading (Barnett, 1986; Carrell, 1988; Reynolds, Taylor, Steffensen, Shirey, & Anderson, 1982; Steffensen, 1987). Bower, Black, and Turner (1979) examined how knowledge of a topic aids in understanding. In this study, students were asked to read a story about a visit to a doctor’s office, and were then asked to recall the material. In 20% of the recalls, students added information that was not in the passage. The authors asserted that because a doctor’s visit is a familiar event (script), the texts were elaborated with background knowledge.

Although it is clear that background knowledge affects reading comprehension strongly, it is also evident that a strong interaction between text processing and the application of background knowledge takes place during the reading process. An example of the strong interaction between text processing at the lowest levels and background knowledge can be seen in the way a reader changes internalized images when reading the following example.

Imagine yourself walking into a room; it is the master bedroom of a quiet Victorian house, in a slum of Bombay, which has just had a fire and been rebuilt in modern style, except for the master bedroom which is only half remodeled having its decorative panelling intact but barely visible because of the thick smoke. (Feldman, 1975, p. 93, presented in Just and Carpenter, 1987)

In this instance, close interaction with the lexical and syntactic strings in the text causes the reader to continually change the content schema being used to produce the mental image. Such interaction of prior knowledge and text processing is an important variable in reading assessment.

Formal Schema. The role of formal schema involves a reader's knowledge of such textual features as coherence, cohesion, and text structure. The focus on these concerns comes from an assumption that there are surface-level phenomena and an internal structure to a text that will help (or hinder) the reader. The surface-level phenomena represent cohesive relations while the internal structures represent the internal structure of the text as a whole.

Cohesive relations are defined as those relations that link one sentence to another without reference to a higher level of analysis. Cohesion occurs when the interpretation of some element in the discourse is dependent on that of another. Grimes (1975) discusses cohesion as the relationship of the new information in the text to information that is already available. Halliday and Hassan (1976) define cohesive ties as instances in which two words are linked by one of five types of relationships: (a) referential; (b) substitution of one word with another; (c) ellipsis; (d) conjunction, additive, adversative, causal, and temporal; and (e) lexical, including reiterations and collocation. Irwin (1986) points out that referential relationships aid text comprehension and that cohesive devices such as conjunctions and connectors help comprehension and reading speed. However, cohesion is different from coherence. Cohesion helps coherence. Van Dijk (1977) asserts that many aspect of language simultaneously contribute to coherence. Connectives, implications, verb frames, property relations, condition-consequence relations, general-particular relations, and other semantic relations that link sentences are all said to add to the coherence of a text. Not surprisingly, studies have shown that readers presented with more and less coherent versions of text show better comprehension in the more coherent versions (McKeown, Beck, Sinatra, & Losterman, 1992). Several areas associated with coherence have shown negative affects on comprehension: (a) ambiguous, distant, or indirect references; (b) indistinct relationships between events; and (c) the inclusion of irrelevant events or ideas (Beck & McKewon, 1989). As such, these areas should be taken into account when selecting texts for assessment and when designing item content.

In addition to cohesive relationships, text structure and organizational pattern play a large role in comprehension. Research has shown that knowledge of text structure interacts with comprehension (Carrell, 1985). For example, text structures with a strong and consistent internal logic are more conducive to processing than are those with weak internal logic (Carrell, 1984). Narratives have a hierarchical schematic structure and when that structure is used to guide comprehension and recall, both are facilitated. Similarly, research with different structures in expository prose has shown effects on reading comprehension due to the location of information within the structures. Specifically, information located high, or at top levels, in the hierarchical organization of a passage is recalled better than

information at lower levels. Tightly organized patterns of comparison, causation, and problem/solution facilitate the recall of specific ideas from a text better than do less constrained text structures such as description (Carrell, 1984). Additionally, Roller (1990) posits an interaction between knowledge and text structure. This line of research has shown that text structure plays a stronger role in comprehension when the content is moderately unfamiliar than when the material is familiar or completely unfamiliar. Given the results of the research on text structure, consideration of the textual organization and the distribution of differing structures across a test should be considered in the test design and construction process.

Strategies and Metacognitive Skills. Readers employ reading strategies and monitor their performance throughout each reading activity. Carrell (1989) asserts that reading strategies are important for what they reveal about the way readers manage their interaction with written text and how these strategies are related to text comprehension. These strategies are employed as the reader regulates comprehension and employs repair strategies. A list of strategies would include such diverse skills as skimming, scanning, guessing meaning in context, skipping unknown words, tolerating ambiguity, reading for meaning, reading critically, making inferences, applying appropriate background knowledge, and recognizing text structure (Carrell, 1991). Barnett (1988) defines two levels of strategy categories, the text level and the word level. Text-level strategies are those strategies related to the reading passage as a whole or to large parts of the passage; they include considering background knowledge, predicting, using titles and illustrations, reading with a purpose, skimming, and scanning. Word-level strategies include such strategies as using context to guess word meanings, identifying the grammatical category of words, following reference words, and recognizing meanings through word families and formation. In terms of comprehension monitoring, Casnave (1988) emphasizes the reader's ability to evaluate the current level of understanding and to plan how to remedy a comprehension problem. Block (1986) identifies main-meaning line strategies and word-solving strategies, and asserts that good readers are more able to monitor their comprehension, are more aware of the strategies they need and use, and use strategies more flexibly than less successful readers. Specifically, good readers adjust their strategies to the type of text they are reading and to their purpose. In the assessment of reading, readers need to be given the opportunity to apply their word-level and text-level strategies in demonstrating their reading ability.

Purpose and Context. From the communicative competence perspective, it is important to address academic reading as a meaning-based activity that is purpose and comprehension driven. Reading is motivated by the reader's particular purpose and is propelled by increasing comprehension of the texts. Further, given that language is embedded in context(s) and tasks, academic reading assessment should focus on contexts and purposes and address narrow language abilities as they emerge from context. Whether we are talking about children reading in school or adults reading university-level course material, it seems clear that most comprehension is linked to purpose, and it is thus important to examine reading within the context of that purpose.

In addressing the context and purpose issues surrounding reading, Hill and Parry (1992) argue against an autonomous model of reading that "treats reading as an autonomous skill that is independent of other factors and transferable across all kinds of texts" (p. 444). Rather, they emphasize the social nature of reading in which reader and writer are involved in an exchange. They assume that readers impute

characteristics to an author, construct some social identity for themselves, and in this process view the text as a form of communication with some social purpose. In this social context, any particular text that is read was written for some purpose, and the reader is likewise reading for some purpose. Consequently, reading is not completely autonomous, but rather is highly contextualized in terms of how and why a reader processes textual information. As Friere (1991) states, "Language and reality are dynamically interconnected. The understanding attained by critical reading of a text implies perceiving the relationship between text and context" (p. 139).

Bennett (1993) has indicated that research on the domain-specific nature of ability has caused cognitive psychology to redirect its orientation in some ways away from general abilities to "situated" cognition. A reader's ability is not a uniform construct in that reading performance will vary depending on such factors as task demand, purpose, and knowledge. Bennett (1993) notes a concomitant need to balance general and context-specific views of ability given that general and specialized knowledge operate together. "Whereas general skills in the absence of a rich domain-specific knowledge base are ineffectual, a domain-specific knowledge base without general skills may only function effectively with formulaic problems" (p. 7-8). Snow & Lohman (1993) indicate that there are sources of variation in both ability and task performance in addition to skill and strategy differences. They note that attentional requirements of tasks, and the degree to which consistent processing becomes automatic with practice, may change the amount of attention the examinee has to devote to the activity. That is, the amount of exposure to the particular task, given the particular purpose, can itself change the examinee's performance. For example, it is likely that in many contexts a reader's comprehension is higher after prolonged exposure to a text or set of texts than it was at the beginning of the text. Thus, assessment needs to account for the broader context of the reading task.

Implications for the Assessment of Academic Reading

Several implications for assessment emerge from the previous discussion of communicative competence and the reading process. These will be discussed below. It should be noted first, however, that in attempting to establish the parameters relevant for development of TOEFL 2000, the TOEFL Committee of Examiners at its spring 1993 meeting, outlined four components of validity that would play central roles in test design. These four components were identified by Messick (1989) as representing a unified validity framework that included the facets of justification of testing and function or outcome of testing. These components are construct validity, value implications, relevance/utility, and social consequences. The following discussion of implications for reading assessment from a perspective of assessing communicative competence will relate implications from the previous discussion on reading to these four components when applicable.

Implication #1. In terms of construct validity, the clearest implication from the previous discussion is the need for expanding beyond the multiple-choice or other selected-response formats. This implication is due to the concerns identified regarding context and purpose, as well as the nature of reading skill application. First, as Mislevy (1993) indicates, tests made up of traditional selected-response items presents tasks to the reader that are more clearly defined and constrained than most real-world problems. Many real-world tasks are ill defined and allow the reader to make choices in precisely how a problem will be solved. Thus, in order to avoid underrepresentation of the construct of academic

reading, constructed-response formats will need to be utilized. Given the discussion of context and purpose, an overreliance on selected-response formats clearly restricts the view of academic reading to an underrepresented construct. For example, the ability to restructure information and apply knowledge is the essence of reading and learning from text. Given that, examinees need to be required to apply knowledge and to restructure the information they confront in a text. It is not possible to do this solely with selected-response formats. A second reason for expanding beyond selected-response formats relates to Bennett's (1993) observation that in some ways tests comprised of selected-response items appear to assume that some skills can be decomposed and isolated from applied contexts. They thus support a view that skills are incrementally acquired. However, the above discussion of the nature of skills points out that skills overlap and are applied in particular contexts differentially depending on, among other things, reading purpose.

The inclusion of constructed-response sections in the test will address some of the negative value implications and social consequences that have become associated with multiple-choice tests. Given that there may be washback effects from assessment onto instruction, the inclusion of more performance-type sections will provide an impetus for teachers and students to focus more on language production and application than on training for the test. Such changes may enhance the relationships between reading and writing as meaning-making activities (Zamel, 1992). However, care will have to be taken to ensure that different subgroups are not adversely affected by the types of tasks and texts that are utilized.

Implication #2. Notwithstanding implication #1, there may still be a need to incorporate multiple-choice or other selected-response formats in order to balance general and context-specific tasks. This inclusion of selected-response items will allow reading assessment to incorporate the way in which both general and specific knowledge interact (Bennett, 1993). Given the importance of automaticity in lexical and syntactic processing, there is a clear need to measure the underlying reading competence of the examinees. It is important to assess both comprehension ability and the application of that comprehension. In this way, traditional tests composed of selected-response items may be viewed as contributing information about the more general cross-contextual components of academic reading. Thus, the test may incorporate a combination of constructed-response and selected-response "items." As Messick (1993) points out, ". . . the question is not simply which format is better for a particular purpose, but also what combination of formats serves that purpose better with the fewest measurement disadvantages" (p. 62). Selected-response formats are efficient mechanisms for establishing the existence and level of underlying enabling abilities.

Implication #3. As noted in the discussion of academic reading, comprehension does not take place in a vacuum. Comprehension in academic settings is translated into some product. The creation of this product involves the application of multiple skills and abilities in carrying out a task or tasks. Thus, reading assessment should be, at least in part, task based (Long & Crookes, 1992). The incorporation of a task basis for assessment will allow the application of many skills on the part of the reader, essentially representing the ecology of reading.

The academic tasks that are used in assessment should encourage the readers to extend themselves and apply reading skills for authentic purposes. Authentic academic tasks, such as answering essay test questions, reading textbooks or journal articles, and taking notes should be identified and incorporated

into the assessment. A representative set of academic reading tasks should be presented to each examinee. This task-based approach recognizes that, generally, the object of comprehension is a portion of a text rather than the complete physical text, and the task being performed generally determines the appropriate portion of text. Further, a focus on tasks will increase the number of contexts that require the reader to apply metacognitive process in monitoring success. By focusing on tasks, then, the test reflects the manner in which reading usually takes place. Likewise, authentic texts should be used, given the importance of text structure and cohesive elements. In order to provide sufficient context for the examinee, the authentic texts should be long enough to provide the full context of the material of interest (Pierce, 1992). However, texts will need to be evaluated for coherence, rhetorical organization, propositional density, and linguistic complexity and so on in order to ensure that the texts examinees confront are cooperative texts. Finally, a focus on authentic tasks and texts will likely have positive value implications and social consequences by emphasizing language in academic context. In terms of value implications, any test-based instruction will need to address the target situations, and in terms of social consequences, the test item types will hopefully not favor any particular group since the tasks are directly reflective of the academic domain. Thus, the tests would reflect tasks that are representative of the ways in which ability is applied in authentic academic contexts, and the extent to which test method is a facet in performance is a direct reflection of academic demands.

Implication #4. As indicated above, background knowledge and context play important roles in reading. Two possible conclusions can be drawn from these observations, given the emphasis that has thus far been placed on purposeful reading. First, although it is important to provide examinees with a variety of texts and text types in order to generalize their abilities beyond a single situation-specific context, it is also important to provide a test that is coherent in terms of its content. This means that the examination may involve thematically based texts in terms of content. Such an organization reflects the large amount of real-world academic reading activity that is organized around particular courses or research activities. The use of thematically related texts addresses Hill and Parry's (1992) observations that in an attempt to present many different kinds of text to an examinee, traditional reading tests typically include several passages on different topics. Thus, each passage is embedded in its own separate textual world, causing a reader to read in a manner much different from authentic reading interactions.

Second, it may be appropriate to develop more adaptive types of tests, either via computer or through selection of texts and tasks by the examinee. This reflects the need for the examination to be coherent in terms of content and from the recognition that much knowledge has a situated basis (Gitomer, 1993). Tests could reflect different content domains, such as natural science, social science, or humanities. Examinees would select the area of interest.

Implications #5. Given that the TOEFL 2000 will be used for academic selection, it is important to consider the relevance and utility components of Messick's (1989) validity model with reference to the new test. A primary element here is the extent to which score interpretation and use may more closely reflect real-world academic reading tasks. Thus, in order to reflect the variability in solving real-world academic tasks, assessment may need to go beyond providing a reading score and present a literacy score that is derived from performance on well defined tasks. Admissions officers may take the applicant's

ability to apply language abilities into account. In order to facilitate this, a descriptive scale related to academic literacy will need to be developed. Furthermore, if score reporting becomes more descriptive and criterion referenced to particular tasks and task domains, rather than strictly norm referenced, the use of the test may conceivably be extended to placement or diagnostic decisions.

Implication #6. From the prior discussion and the implications drawn above, additional general recommendations about test design seem warranted. First, in principal, reading could be integrated with any of the other language skills. However, this should be based on the “literacy” task being encountered. The test format could be such that the examinee acquires information on a particular topic from reading and/or listening, and then produces some written or oral product. For example, the examinee might read a passage and summarize the information. Summary tasks can represent what the examinee understands about a text as he or she reads. A written summary can indicate the gist in the examinee’s mind, and this may represent overall comprehension of a passage (Kintsch & van Dijk, 1978). The writing summary might be followed by a listening passage on a similar topic. The examinee might then answer selected-response items related to information presented in the passage and then write an evaluative essay. Score reporting could then indicate both a reading score and a functional academic literacy score.

Conclusions

The preceding discussion presents an initial view of how to communicatively assess academic reading ability. The approach described viewing reading in a literacy context. It stresses moving beyond seeing reading as an autonomous ability that is assumed to be relatively static across texts and tasks. Such a view involves regarding reading as a purposeful activity within a context. Additional emphasis is given to the need to incorporate constructed-response formats into the assessment in order to address the constructive and reconstructive aspects of reading in context. Finally, this discussion recommends that reading assessment be integrated with other language skills.

In proposing the above, there is no assertion that this will be easy. Several issues remain to be resolved. Some of these issues are:

1. How will scores be reported? Will they present a continuous scale or represent threshold behaviors?
2. How will test security be maintained?
3. How will equivalence be ensured?
4. How will task fairness be assured?
5. How systematic will the integration of skill areas have to be across administrations?
6. How will reliability be ensured?
7. How will any negative effects due to subjectivity of evaluation be minimized?

This is certainly not an exhaustive list of questions that remain to be answered. It indicates, however, many of the issues that will arise as the TOEFL 2000 is developed to reflect the communicative nature of academic reading.

References

- Alderson, J. C. (1990). Testing reading comprehension skills (part one). *Reading in a Foreign Language*, 6:425-438.
- Anderson, R. C., & Pearson, P. D. (1988). A schema-theoretic view of basic processes in reading comprehension. In P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive Approaches to Second Language Reading* (pp. 37-55). Cambridge: Cambridge University Press.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L., & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *The Modern Language Journal*, 70, 380-390.
- Balota, D. (1990). The role of meaning in word recognition. In D. Balota, G. d'Arcias, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9-32). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Barnett, M. A. (1986). Syntactic and lexical/semantic skill in foreign language reading: Importance and interaction. *The Modern Language Journal*, 70, 343-349.
- Barnett, M. A. (1988). Reading through context: How real and perceived strategy use affects L2 comprehension. *The Modern Language Journal*, 72, 150-162.
- Beck, I. L., & Carpenter, P. A. (1986). Cognitive approaches to understanding reading: Implications for instructional practice. *American Psychologist*, 41(10), 1098-1105.
- Beck, I., & McKeown, M. (1989). Expository text for young readers: The issue of coherence. In L. Resnick (Ed.), *Knowing, learning, and instruction* (pp. 47-65). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bennet, R. (1993). On the meanings of constructed response. In R. Bennett & W. Ward (Eds.), *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (pp.1-27). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, 20:463-494.
- Bower, G., Black, J., & Turner, T. (1979). Scripts in memory for text. *Cognitive Psychology*, 11, 177-220.
- Campbell, R., & Wales, R. (1970). The study of language acquisition. In J. Lyons (Ed.), *New horizons in linguistics*. Harmondsworth: Penguin Books.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 147.

-
- Carrell, P. (1984). The effects of rhetorical organization on ESL readers. *TESOL Quarterly*, 17, 441-469.
- Carrell, P. L. (1985). Facilitating ESL reading by teaching text structure. *TESOL Quarterly*, 19, 727-752.
- Carrell, P. L. (1988). Some causes of text-boundedness and schema inference in ESL reading. In P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 101-113). Cambridge: Cambridge University Press.
- Carrell, P. L. (1989). Metacognitive awareness and second language reading. *The Modern Language Journal*, 73, 121-134.
- Carrell, P. L. (1991). Strategic reading. In: Georgetown University Round Table on Languages and Linguistics (pp. 167-178). Washington, DC: Georgetown University Press.
- Casnave, C. P. (1988). Comprehension monitoring in ESL reading: A neglected essential. *TESOL Quarterly*, 22(2), 283-302.
- Cunningham, A., Stanovich, K., & Wilson, M. (1990). Cognitive variation in adult college students differing in reading ability. In T. Carr & B. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 129-159). San Diego, CA: Academic Press.
- Feldman, J. (1975). Bad-mouthing frames. In R. Schank & B. L. Nash-Webber (Eds.), *Tinlap conference proceedings*. Cambridge MA: The MIT Press.
- Friere, P. (1991). The importance of the act of reading. In C. Mitchell & K. Weiler (Eds.), *Rewriting literacy: Culture and the discourse of the other* (pp. 139-145). New York: Bergin & Garvey Publishers.
- Gick, M., & Holyoak, K. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.
- Gitomer, D. (1993). Performance assessment and educational measurement. In R. Bennet & W. Ward (Eds.), *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (pp. 241-263). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Goodman, K. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 6, 126-135.
- Gough, P. B. (1972). One second of reading. In J. F. Kavanaugh & I. G. Mattingly (Ed.), *Language by eye and ear* (pp. 331-358). Cambridge: MIT Press.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, 25, 375-406.

-
- Green, J. (1987). In search of meaning: A sociolinguistic perspective on lesson construction and reading. In D. Bloome (Ed.), *Literacy and Schooling* (pp. 4-34). Norwood, NJ: Ablex Publishing Corp.
- Grimes, J. (1975). *The Thread of Discourse*. The Hague, Netherlands: Mouton.
- Halliday, M., & Hassan, R. (1976). *Cohesion in English*. London: Longman.
- Haynes, M., & Carr, T. (1990). Writing system background and second language reading: A component skills analysis of English reading by native speaker-readers of Chinese. In T. Carr & B. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 375-421). San Diego, CA: Academic Press.
- Henk, W., & Helfeldt, J. (1987). In-process measures of ambiguous text interpretation: Another look at the influence of prior knowledge. In J. Readence & R. Baldwin (Eds.), *Research in literacy: Merging perspectives: 36th Yearbook of the National Reading Conference* (pp. 143-149). Rochester: National Reading Conference, Inc.
- Hill, C., & Parry, K. (1992). The test at the gate: Models of literacy in reading assessment. *TESOL Quarterly*, 26, 433-461.
- Hudson, T. (1991). A content-comprehension approach to reading English for science and technology. *TESOL Quarterly*, 25, 77-104.
- Hudson, T. (1993). Testing the specificity of ESP reading skills. In D. Douglas & C. Chapelle (Eds.), *A New Decade of Language Testing Research* (pp. 58-82). Alexandria, VA: TESOL.
- Hymes, D. (1972). On communicative competence. In J. Pride & J. Holmes (Eds.), *Sociolinguistics*. Harmondsworth: Penguin.
- Irwin, J. W. (1986). Cohesion and comprehension: A research review. In J. W. Irwin (Ed.), *Understanding and teaching cohesion comprehension* (pp. 31-34). Newark, Delaware: International Reading Association.
- Johnston, P. (1984). Prior knowledge and reading comprehension test bias. *Reading Research Quarterly*, 19, 219-239.
- Just, M., & Carpenter, P. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.
- Kintsch, W., & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.

-
- Long, M., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, 26, 27-56.
- Marshall, S. (1993). Assessing schema knowledge. In N. Frederiksen, R. Mislevy, & I. Bejar (Eds.), *Test theory for a new generation of tests* (pp. 150-180). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McKeown, M. G., Beck, I. L., Sinatra, G. M., & Losterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension. *Reading Research Quarterly*, 27, 79-93.
- Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational Measurement (3rd Edition)* (pp. 13-103). New York: American Council on Education — Macmillan Publishing Company.
- Messick, S. (1993). Trait equivalence as construct validity of score interpretation across multiple methods of measurement. In R. Bennet & W. Ward (Eds.), *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, portfolio assessment* (pp. 61-73). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mislevy, R. (1993). Foundations of a new test theory. In N. Frederiksen, R. Mislevy, & I. Bejar (Eds.), *Test theory for a new generation of tests* (pp. 19-39). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Perfetti, C. (1992). The representation problem in reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 145-174). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Peirce, B. (1992). Demystifying the TOEFL reading test. *TESOL Quarterly*, 26, 665-691.
- Pollatsek, A., & Rayner, K. (1990). Lexical access: Some comments on models and metaphors. In D. Balota, G. d'Arcias, K. Rayner (Eds.), *Comprehension processes in reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rayner, K. (1990). Comprehension processes: Introduction. In D. Balota, G. d'Arcias, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 1-6). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs: Prentice Hall.
- Reynolds, R. E., Taylor, M. A., Steffensen, M. S., Shirey, L. L., Anderson, R. C. (1982). Cultural schemata and reading comprehension. *Reading Research Quarterly*, 17, 353-366.
- Roller, C. (1990). Commentary: The interaction of knowledge and structure variables in the processing of expository prose. *Reading Research Quarterly*, 25, 79-89.
- Rumelhart, D. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance IV* (pp. 573-603). New York: Academic Press.
- Savignon, S. (1991). Communicative language teaching: State of the art. *TESOL Quarterly*, 25, 261-277.

-
- Sharkey, N. (1990). A connectionist model of text comprehension. In D. Balota, G. d'Arcias, & K. Rayner (Eds.), *Comprehension Processes in Reading* (pp. 487-514). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Skehen, P. (1991). Progress in language testing: The 1990s. In J. Alderson & B. North (Eds.), *Language testing in the 1990s: The communicative legacy*. London: Macmillan Publishers, Ltd.
- Smith, F. (1971) *Understanding Reading*. New York: Holt, Rinehart, & Winston.
- Snow, R., & Lohman, D. (1993). Cognitive psychology, new test design, and new test theory: An introduction. In N. Frederiksen, R. Mislevy, & I. Bejar (Eds.), *Test Theory For A New Generation of Tests* (pp. 1-17). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.
- Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32-71.
- Steffensen, M. S. (1987). The effect of context and culture on children's L2 reading: A review. In J. Devine, P. Carrell, & D. E. Eskey (Eds.), *Research in reading in English as a second language* (pp. 43-54). Washington: TESOL.
- Symons, S., & Pressley, M. (1993). Prior knowledge affects text search success and extraction of information. *Reading Research Quarterly*, 28, 251-259.
- van Dijk, T. (1977). *Text and context*. London: Longman.
- van Dijk, T. (1985). Strategic discourse comprehension. In T. Ballmer (Ed.), *Linguistic dynamics: Discourses, procedures and evolution* (pp. 30-61). Berlin: Walter de Gruyter & Co.
- Valencia, S. W., & Stallman, A. (1989). Multiple measures of prior knowledge: Comparative predictive validity. *Yearbook of the National Reading Conference*, 38, 427-436.
- Zamel, V. (1992). Writing one's way into reading. *TESOL Quarterly*, 26, 463-485.



Cover Printed on Recycled Paper

58701-13990 • Y96M.75 • 253701 • Printed in U.S.A.