

Cognition and Reading: Cognitive Levels as Embodied in Test Questions

J. Charles Alderson and Yasmeen Lukmani

University of Lancaster and Bombay University

Various lists of reading skills are proposed in the literature. It has frequently been asserted, or implied, that these skills can be ranked hierarchically, from 'lower' to 'higher' order skills. In this study, teachers at Lancaster University were presented with a reading test used at Bombay University and required, among other tasks, to rate the items as tests of 'lower', 'middle', or 'higher' order abilities. In 27 out of 41 items, there was little agreement between judges. An item analysis of the performance of students at Bombay University was carried out on the remaining 14 items, on which agreement had been reached. This showed that there was little relationship between level of item difficulty and the supposed level of comprehension. Discrimination was largely achieved by the lower order items. Some possible implications for the testing of language and reading are put forward.

INTRODUCTION

It is commonplace in the teaching and researching of reading – in mother tongue as well as in a foreign language – to distinguish different levels of understanding of a given text. The distinction between “literal understanding” and “inferencing meanings that are not directly stated in a text” is fairly familiar to teachers of reading, as is Gray’s (1960) distinction between “reading the lines”, “reading between the lines” and “reading beyond the lines”. This distinction can be illustrated by the following examples.

1. Twas brillig and the slithy toves
Did gyre and gimbol in the wabe.
All mimsy were the borogroves
And the mome raths outgrabe

Although readers of “Jabberwocky” may say they do not understand the poem, it is clearly possible to answer questions about it such as:

How were the toves?	Answer . . . slithy
What did the toves do?	Answer . . . gyre and gimbol
Where did the toves gyre and gimbol?	Answer . . . in the wabe

Readers use their knowledge of English syntax and morphology to arrive at answers to the questions, and at an understanding, at some “level” of the poem. Most people would agree, however, that their “understanding” of the poem is

J. Charles ALDERSON is a Senior Lecturer in the Department of Linguistics, University of Lancaster. His main interests are language testing and reading research. Yasmeen LUKMANI is Reader in the English Department, Bombay University. She took her Masters degree at UCLA, and has published on motivation in EFL learning. The authors may be contacted via J. Charles Alderson, Department of Linguistics, University of Lancaster, Lancaster.

incomplete, despite their ability to answer such questions.

Even if one understands the words in sentences, comprehension is commonly held to go beyond the ability to understand the “literal meaning” of words and sentences. Inferencing is an essential and common feature of text processing. Consider the sentences presented in Bransford et al (1984):

- 2a) The floor was dirty so Sally used the mop.
- 2b) The floor was dirty because Sally used the mop.

Experimental subjects reported having seen the sentence “the mop was dirty” in recall tests after sentence 2b) had been presented to them, but not after having read sentence 2a). Since the sentence “the mop was dirty” had not been presented to them, they must have inferred it from a processing of sentence 2b).

Consider next the following text, also from Bransford et al:

- 3) Sally first tried setting loose a team of gophers. The plan backfired when a dog chased them away. She then entertained a group of teenagers and was delighted when they brought their motorcycles. Unfortunately, she failed to find a Peeping Tom listed in the Yellow Pages. Furthermore, her stereo system was not loud enough. The crab grass might have worked but she didn't have a fan that was sufficiently powerful. The obscene phone calls gave her hope until the number was changed. She thought about calling a door-to-door salesman but decided to hang up a clothesline instead. It was the installation of blinking neon lights across the street that did the trick. She eventually framed the ad from the classified section.

In an experiment to test comprehension of this and similar passages, Bransford et al report that subjects have considerable difficulty in understanding this passage until they are given a suitable title (“Getting Her Neighbours To Move”). Subjects were able to understand the individual sentences, but could not make them cohere into a larger text, because they did not know what the passage was about.

Given that these examples evidence the possibility of understanding at different levels – that is the level of understanding the words, or the level of understanding the syntax, or the level of understanding the sentences, or the level of understanding the connection between sentences – and given that teachers themselves are quite accustomed to the notion of talking about students' ability to understand the sentences or the details of a text, but not to grasp the main idea, it seems at least intuitively worthwhile to make distinctions between and among levels of comprehension, and this indeed is what traditional reading research has done.

PREVIOUS RESEARCH

Reading researchers have long attempted to identify reading skills or subskills by giving subjects a series of passages and asking them questions intended to test different levels of understanding of the passages. The answers to these questions are then subjected to factor analysis, in order to see whether the different questions measure different “subskills”. Many different lists of subskills have been drawn up as a result of research like this, from the thirty-six skills identified by the New York City Board of Education (mentioned in Lunzer and Gardner 1979) to the three identified by Gray (1970) quoted above. Davis (1944) concluded that there were nine distinguishable skills in reading, although his evidence was challenged by later researchers (Thorstone 1946). Davis (1968) defined eight skills and devised items to test them. These were:

- 1) Recalling word meanings
- 2) Drawing inferences about the meaning of a word from context
- 3) Finding answers to questions answered explicitly or in paraphrase
- 4) Weaving together ideas in the content
- 5) Drawing inferences from the content
- 6) Recognizing a writer’s purpose, attitude, tone and mood
- 7) Identifying a writer’s technique
- 8) Following the structure of a passage

It should be fairly clear from these descriptions of the “subskills” that they relate directly to the level of understanding that a reader has of a text or texts.

Argument has raged over the years as to whether these “skills” are separately identifiable, or whether they relate to common underlying abilities. Davis himself produced evidence that skills 1, 3, 5, 6 and 8 above are experimentally distinguishable. However, Thorndike (1971) reanalysed the data and claimed that only the first skill – word knowledge – could be distinguished from the others by factor analysis, and concluded that there was a large first factor in tests of reading comprehension, rather than separable skills relating to levels of understanding. Subsequent reanalysis of the same data by Spearitt (1972) apparently identified four separate factors: recalling word meanings; drawing inferences from the content; recognizing a writer’s purpose, attitude, tone and mood; following the structure of a passage; yet concluded that “present types of reading tests, as distinct from word knowledge tests, largely measure one basic ability, which may well correspond to the label of ‘reasoning in reading’” (page 110).

Other researchers have distinguished “comprehension skills” from “inference skills” (Carrol 1969, 1971), thereby acknowledging differences in what is understood from texts. Munby (1978) has devised a list of “microskills” that he considers to contribute to readers’ abilities to understand texts, and his list – see Table One

below – has had enormous influence in the area of the teaching and testing of English as a Foreign Language. It should be emphasised, however, that Munby arrived at his list of “skills”, not by devising questions to measure different subskills, and then subjecting them to empirical verification, but by thinking about the nature of language and speculating – some might say introspecting – on what a reader must logically have to do in order to understand written texts. Again, however, these skills must relate in some way to the meaning that a reader derives from texts, that is, to the nature – or “level” – of understanding of texts.

Table 1: Reading ‘Microskills’, from Munby (1978)

- Recognizing the script of a language
- Deducing the meaning and use of unfamiliar lexical items
- Understanding explicitly stated information
- Understanding information when not explicitly stated
- Understanding conceptual meaning
- Understanding the communicative value of sentences
- Understanding relations within the sentence
- Understanding relations between parts of text through lexical cohesion devices
- Understanding cohesion between parts of a text through grammatical cohesion devices
- Interpreting text by going outside it
- Recognizing indicators in discourse
- Identifying the main point or important information in discourse
- Distinguishing the main idea from supporting details
- Extracting salient points to summarize (the text, an idea)
- Selective extraction of relevant points from a text
- Basic reference skills
- Skimming
- Scanning to locate specifically required information
- Transcoding information to diagrammatic display

Other writers have been influenced by the sort of taxonomies drawn up by Benjamin Bloom and his associates (Bloom 1956), who attempted to define educational objectives in the cognitive domain, and who by extension arrived at a definition of cognitive skills that has had considerable influence on those who think about psychological abilities and psycholinguistic processing.

Bloom’s taxonomy consists of the following educational objectives: Knowledge, Comprehension, Application, Analysis, Synthesis, Evaluation. ‘Comprehension’, in turn, is sub-divided into Translation, Interpretation, and Extrapolation.

In the world of the teaching of English as a Foreign Language, one attempt to apply

Bloomian taxonomies to the teaching and testing of reading is illustrated by Adams-Smith (1981). Readers are invited to read through the text below presented by Adams-Smith, to attempt to answer the questions that follow and to ask themselves what sorts of things they have to do with the text in order to arrive at satisfactory answers to the questions.

HEAT LOSS FROM THE HUMAN BODY

Heat loss by sweating depends on the fact that when a liquid evaporates, it absorbs an enormous quantity of heat from its surroundings. In the case of water, this amounts to 2256j for every millilitre that evaporates. Therefore, when 1ml. of sweat evaporates, a great deal of heat is absorbed from the surface of the body in contact with it. This heat transfer occurs even if the environment is hotter than the body.

Two factors affect the rate of evaporation of sweat, and therefore the effectiveness of sweating as a method of cooling the body. The first is the amount of movement of air surrounding the body. The second is the amount of water vapour in the air which surrounds the body.

When air moves over the surface of water, the amount of evaporation is greatly increased. This is why housewives prefer to hang out their washing on windy days: it dries more quickly than on still days. For the same reason, sweat evaporates very rapidly on windy days, and the rate of heat loss by sweating is much more than on a still day. This accounts for the fact that hot, still days are much less comfortable than hot, windy days. In contrast, the sweat evaporates very rapidly on a hot, windy day, and cools the body quickly and effectively.

The second factor that affects the rate of evaporation of sweat is the amount of water vapour in the air, that is, the humidity. When air is carrying the maximum amount of water vapour that it can hold, it is said to be 100 per cent saturated with water vapour. The relative humidity of the air is said to be 100 per cent. Under these conditions the air cannot carry any more water, so no water can evaporate. When the relative humidity is high, therefore, sweat cannot evaporate. Instead, it forms large drops and runs off your skin without cooling you.

When the air is very dry and carries no water at all, the relative humidity is said to be 0%. It is obvious that under these conditions evaporation will be much more rapid. Therefore, sweating will be much more effective as a method of losing heat from the body. On a hot, dry day sweat evaporates as soon as it is formed, and you feel reasonably cool even though the temperature of your environment is very high.

- Memory** without referring back to the passage, answer the following question:
What are the factors that affect the rate of evaporation?
- Translation:** Draw a simple graph to show the relation between the amount of air movement, and the rate of sweat evaporation.
- Interpretation:** a) What is the relationship between heat loss and environmental temperature?
b) Compare Figure 1, and the graph you drew in answer to question 2.
- Application:** During which of the following weather conditions would you expect (i) the highest incidence, (ii) the lowest incidence of heat stroke? Give reasons.

<i>Max. temperature</i>	<i>Humidity</i>	<i>Wind Velocity</i>
105	92%	zero
95	10%	15 mph
100	50%	5 mph
110	0%	20 mph

- Analysis:** a) Explain how heat transfer occurs when the environmental temperature exceeds 99° F.
b) Analyse the reasoning in the following statement: "You do not need to drink so much when it is dry, as the moisture in the air helps to keep you cool".
- Synthesis:** Design a heat stroke trauma unit for a Kuwait hospital.
- Evaluation:** a) Decide which climate is more dangerous to health: that of Kuwait in summer, or of Alaska in winter.
b) Evaluate the care received by heat-stroke victims in Kuwait, and suggest measures for improving it.¹

'LEVELS' OF SKILLS

When discussing the skills or subskills that might be involved in reading, it is fairly common to make a distinction between "higher" and "lower" order skills. What are typically considered to be "lower-order" skills are those mentioned first in most lists of skills. Thus, in the Adams-Smith list above, memory, translation and

¹Adam-Smith's 'Memory' seems to be equivalent to Bloom's 'Knowledge'. Bloom's category 'Extrapolation' is missing from Adam-Smith's illustration; otherwise her scheme follows Bloom very closely.

interpretation are typically held to be “lower” than analysis, synthesis and evaluation, and Davis’ skills of recalling word meaning and finding answers to questions answered explicitly are “lower” than drawing inferences or recognising writer’s purpose or tone.

Although this distinction is often asserted or assumed, there is relatively little research that has attempted to establish empirically that such a distinction is justified. Lunzer, Waite and Dolan (1979) attempted to devise reading comprehension tests aimed at assessing English native speakers’ abilities to understand text at different levels of comprehension, through questions intended to tap different reading skills. The skills intended to be tested were:

- Word meaning
- Words in context
- Literal comprehension
- Drawing inferences from single strings (sentences)
- Drawing inferences from multiple strings (sentences)
- Interpretation of metaphor
- Finding salients or main ideas
- Forming judgments

The authors failed, however, to find evidence for the separability of such skills, and in particular were unable to demonstrate that these skills were arranged in a hierarchy on an implicational scale. It is sometimes asserted that reading subskills relate to each other in such a way that before one can employ higher-order skills, one has to master preceding lower-order skills, yet Lunzer et al. were unable to identify readers who were, for example, able to answer word-meaning questions but not questions further up the supposed hierarchy, or readers who were able to understand the literal meaning of sentences but could not draw inferences, or readers who could understand the main idea of a passage yet not evaluate its significance. In fact, subjects proved able to answer “higher-order” questions yet failed to answer “lower-order” questions correctly.

THE PRESENT STUDY

These findings have upset quite a few theorists and approaches to the study of reading and, therefore, it seemed important to attempt to replicate them with non-native speaker populations, since the general distinction between lower and higher order skills is particularly influential in the teaching of English as a Second or Foreign Language. What follows is a report of research conducted at Bombay University, India, and Lancaster University, England.

The research was directed at a reading test used at the University of Bombay. This test, the Final Year BA Communications Skills in English Examination, is sat by

students at the end of their first year of undergraduate study. They have completed a compulsory course in Language & Communication Skills as a complement to their academic studies, these studies being in a variety of subjects. The Communication Skills Course, in other words, is a form of ESP teaching. Students have to pass a test at the end of the year in order to show evidence of their ability to read in English. It is the belief of the exam writers and the course designers, that it is important to test higher order skills of comprehension as well as lower order skills, and that it is important to teach these as well. The skills being tested by these reading comprehension questions, as described by the Bombay University Communication Skills Group, are listed in Table Two below.

Table 2: Skills being tested (possibly) by reading comprehension questions

1. **Recognition of words** (*Vocabulary*)
 - i) *Lexical matching of words/phrases similar in meaning or opposed in meaning (reference to text is incorporated in the question)*
 - ii) *Relating of given non-textual word/phrase to word/phrase contained in the text (reference to text required of student)*
2. **Identification:** *Ability to locate and recognise information in one spot/in more than one spot with or without rephrasing within the specific context*
3. **Discrimination:** *Ability to distinguish/select between elements/objects/features within the specific context in which they occur.*
4. **Analysis:** *Ability to break down entities into their constituent parts, to perceive the essential features of entities and their relations, to examine methodically:*
 - i) *the constituent elements within a structure*
 - ii) *the relationship between elements, e.g. causal/sequential/chronological/hierarchical/ordering of elements*
 - iii) *the logic of an argument*
5. **Interpretation:** *Ability to clarify the meaning of complex ideas or configurations and to interpret the significance of ideas/actions/events/relationships*
6. **Inference:** *Ability*
 - i) *to derive conclusions from facts and premises*
 - ii) *to predict the continuation of arguments/events*
7. **Synthesis:** *Ability to synthesize, to recognise form, pattern, attitude*
8. **Evaluation:** *Ability to appraise, evaluate, judge.*

The research was intended to investigate the question "Do identifiable separate levels of comprehension exist?" We attempted to answer the question in two ways,

- (a) judgementally, using 9 teachers at the University of Lancaster, and (b) empirically, using 100 students at Bombay University.

JUDGEMENTAL PHASE

In this phase of the study, 9 teachers in the Institute for English Language Education in the University of Lancaster were given a task which is set out below. The task, in effect, involves making judgements about what is being tested by each of the test items in the Bombay University test.

Task

1. Please read through the accompanying test paper (*F Y B A Communications Skills in English*) (*Final Examination*) and answer each question to the best of your ability, as if you were indeed taking a reading test
2. On a separate piece of paper, preferably with your name on, please write down for each of the three passages, what **you** think each test question and each sub-question is testing. Please use your own concepts, vocabulary, views of what reading might be and what test content might be, but please try to be explicit (i.e. so that someone not necessarily sharing your own jargon, concepts or view of the world could (hopefully) understand).

THIS WILL TAKE SOME TIME.

IF POSSIBLE, REFER BACK TO WHAT YOU HAD TO DO IN ORDER TO ANSWER EACH QUESTION

3. Now, classify each question or sub-question according to whether it could be said to measure

<i>LOWER</i>	<i>Order</i>	
<i>MIDDLE</i>	<i>Order</i>	<i>abilities. These are undefined</i>
<i>HIGHER</i>	<i>Order</i>	

4. Finally – **AND PLEASE DO NOT DO THIS UNTIL YOU HAVE FINISHED THE ABOVE THREE STEPS** – on a separate piece of paper, please indicate for each question, which of the skills on the piece of paper in the accompanying envelope is/are being tested by each question or sub-question. You may, if you wish, choose more than one of these for any question, but please do not add other skills.

PLEASE DO NOT GO BACK AND CHANGE YOUR ANSWERS TO THE FIRST THREE TASKS ABOVE!!!!

The envelope contained the skills supposedly being tested by the reading test, exactly as set out in Table 2 above.

Having obtained the responses of the 9 "expert" judges, it was possible to pool the both on the level of skill being tested by any one item, and on what each item was thought to be testing. Results were perhaps surprising. There was considerable disagreement amongst judges as to the level of the question and the content of the item. For example, in one question one judge said the item was a test of lower order skills, a second judge said the item was a test of middle order skills and a third judge said the item was a test of higher order skills. For example, see item 2. liv in Table 4 below.

Table 4: Level of item (High, Mid, Low) according to Judge Number

Item	1	2	3	4	5	6	7	8	9
1.1	L	H+L	M	L	L	L	L	L	M
1.2	L	L	M	L	L	L	L	L	M
1.3	M	L	M	L	L	L	L	M	H
1.4a	M	L	M	M	M	L	M	M	M
1.4b	L	L	M	M	M	L+	M	M	M
1.5	M	L	M	L	L	L+	?	L	H
1.6	L	L+M	M	L	L	L	?	L	M
1.7	L	L+M	M	L	M	?	M	L	M
1.8	L	L	M	L	L	?	L	L	M
1.9	L	L	M	L	L	L+	L	L	M
2. li	H	H	?H	M	H	M	M+H	M	H
2. lii	H	H+L	?H	M+H	H	M	M/H	L	H
2. liii	M	L/M	H	H	H	M	M+H	L	H
2. liv	H	L	?H	M	H	?	M/H	L	H
2. lv	H	L	?H	M	H	?	M/H	L	H
2. vi	H	H	?H	—	H	?	M/H	L+	H
2. lvii	—	L+H	?H	H	H	L+	M/H	L	H
2. lviii	H	H+M	?H	H	H	M	M/H	L	H
2. lix	H	H,M	?H	—	H	L+	M/H	L	H
2. lx	M	L	?H	N	H	L+	M/H	L	H
2. lxi	H	N	?H	H	H	L	H		
2. lxii	H	H	?H	M	H	L+	M/H	L	H
2. lxiii	M	L	?H	M	H	L	M/H	L+	H
2. lxiv	M	??	?H	M	H	M	M/H	L	H
2. lxv	—	L	?H	M	H	?	M/H	L+H	
2. lxvi	—	L	?H	M	H	?	M/H	L	H
2.2	H	M+H	H	H	?	L	H	H	
2.3	H	M+H	H	H	H	M	M	H	H

2.4	H	H+M	?H	H	M/H	??	H	H	H
2.5	H	M+H	H	H	M/H	?	M	H	H
2.6	H	H	H	H	H	?	H	H	H
3.1	L	LHM	H	H	M/H	H	H	L+	M
3.2	M	L+M	H	M	H	H	H	M	M
3.3	M	M	M	M	M	?	M?	L	M
3.4	H	L+M	H	H	H	M+	M?	M+	H
3.5	M	M	M	M	M	?	—	L+	H
3.6	H	M	H	H	M/H	M	—	H	H
3.7	L	L	M	M	H	M	—	L	L
3.8	L	L?M	M	M	H	M/H	—	L	H
3.9	M	L	M	M	H	M/H	—	H	L
3.10	L	L	H	M	M	M/H	—	L	M
3.11	H	H	H	H	H	?	—	H	H

RESULTS

For 27 out of the 41 items there was very little agreement on the levels being tested. In addition, for approximately the same number of items judges disagreed considerably over the skills being tested. So, for example, taking the list of skills as detailed above in Table Two, some judges regarded one question as testing level one, others as testing level three, others testing level four, others level five, others level seven and so on.

When asked to describe what each item was testing, the judges sometimes gave very different accounts of what was involved. For example, the first test question reads:

1.1 Where does the author see the sights he has described in this passage?

The descriptions given by each judge were as follows

- Judge 1: Recognition of vocabulary and grammar
- Judge 2: The words “where”, “bottom”, “Atlantic”; knowledge that settings are early in a narrative
- Judge 3: Scan for specifics, relations in a sentence
- Judge 4: Detailed information, recognise prepositions, links between parts of complex
- Judge 5: Inferred meaning; discourse rules re mentions
- Judge 6: Direct verbatim extraction of information
- Judge 7: Establish context, lift a phrase direct from text
- Judge 8: Direct reference. intersentence connections
- Judge 9: Understand text links

Another example is question 2.4

- 2.4 The reference to the noble savage in paragraph 2 is meant to
- A) defend
 - B) illustrate
 - C) explain
 - D) contradict

The descriptions of the judges were as follows:

- Judge 1: Recognition of the tools of an argument
- Judge 2: Function of 2nd paragraph, understanding main point
- Judge 3: Explicitly stated information, gist
- Judge 4: Recognise discursal development and role of paragraph in this; implicit knowledge
- Judge 5: Knowledge of literary style; distinguish related terms
- Judge 6: Awareness of textual relations and functions
- Judge 7: Recognise cues to cultural, educational background
- Judge 8: Relate paragraphs to each other, identify main idea
- Judge 9: Understand coherence, pick out main ideas

In fact, it looks as if expert judges are not able to agree on the skills and levels being tested by any particular comprehension test item over half the time. If so-called expert judges are not able, with confidence, to identify level or skill of comprehension, then one must surely begin to question whether definably separate or different levels do exist.

It must be said that there is likely to be a considerable degree of unreliability in the judgements made. Although we have no data, it is suspected that if the 10 judges were asked to re-judge the questions they have already judged one week later, there would be considerable variation – at least in the skills being identified on each question. Without counter evidence, this must remain a problem.

In addition, it is likely that the manner in which individuals arrive at any answer to a question will vary. One person may have difficulty with a particular word and need to infer connections across sentences. Another may understand the word and, therefore, not need to infer. Thus a product, that is, a right answer, may be arrived at in a variety of different ways using different processes, strategies or skills. In other words the product does not equal the process. This is likely to be true for judges as well as for individual test-takers.

In reading research it is common to infer skills and processes from answers to test questions. It is important to remember that such an activity is a high inference activity. It may not be legitimate to infer a process or skill from a product or answer. Thus, we have to be cautious in believing that certain processes do indeed underlie particular responses. Recently, reading research has seen the beginnings of a

tendency to explore with students how they understand a text rather than to rely on an observer's inferences. Talking to students about how they have arrived at particular answers, what makes them think they are right, what exactly they have understood, above all getting students to justify and explain their answers, should enable us to understand their understanding. A fruitful avenue for further research into skills and levels of comprehension would be to carry out such a de-briefing with students after they have taken comprehension tests, in order to understand understanding.

Nevertheless, be that as it may, in the case of the present study, considerable agreement among judges was indeed achieved in 14 out of the 41 reading comprehension questions. That is, judges largely agreed with each other on whether the item was testing lower, middle or higher order skills and a degree of agreement, although by no means unanimity, was achieved on the skills being tested by these 14 items. So for something less than half the items involved, it is possible to say that some identification of level seems to be possible judgementally.

It was decided, therefore, in the empirical phase of the study to proceed with further investigation only of those 14 items which judges had, more or less, agreed upon.

EMPIRICAL PHASE

In this phase of the study, the results of the performance of 100 students at Bombay on the above-mentioned 14 items were subjected to item analysis. These items were as follows:

LOWER

- 1.2 Who is his companion?
- 1.5 Which expression indicates that the author was frightened by what he saw?
- 1-6 What were the innumerable shining points that the author could see in the darkness?
- 1.8 The author uses different adjectives which indicate a very big size. Mention any two such adjectives.
- 1.9 Match the words given under List A with the nearest meaning (as used in the passage) given under List B. List B has some extra items.

List A

- a) profoundly
- b) irradiation
- c) impenetrable
- d) cavity

List B

- i) illumination
- ii) inaccessible
- iii) extremely
- iv) densely
- v) inescapable
- vi) hollow

MIDDLE

- 1.4a) How long did it take them to come near the spot which was brightly lit up?
- 1.4b) How much distance did they need to cover in order to approach it?
- 2.1.iii Doctors are unsure about the overall implications of diet.
True or False?
- 3.3 "It seizes with avidity upon any incidents . . ." What does 'it' refer to in this sentence?
- 3.5 "It is the protest of romance against the commonplace of life".
The protest refers to
- A) the invention of a legend
 - B) a fanatical belief
 - C) a passport to immortality
 - D) the philosopher's smile

HIGHER

- 2.1.viii The main causes of disease are sugar, fats, butter and eggs.
True or False?
- 2.1.xi The coarse bread eaten by the Victorians was healthier than our refined bread. True or False?
- 2.2 This passage can be called
- A) critical
 - B) analytical
 - C) ironical
 - D) cynical
- 2.3 The passage seems to be written for
- A) a group of dieticians
 - B) food processors
 - C) the common reader
 - D) food technologists
- 3.11 This passage can best be described as
- A) reflective
 - B) descriptive
 - C) informative
 - D) subjective

It can be seen that of these questions, there were 5 lower order items, 4 middle order items and 5 higher order items. The results of an item analysis on the performance of 100 Bombay University students on these 14 items are contained in Table Five below.

Table 5: Results of item analysis on Bombay University Reading Test

Test Item	Top Group n=50	Bottom Group n=50	Total %	Discrimination E 1-3
LOWER				
1.2	49	40	89	.18
1.5	49	23	72	.52
1.6	48	14	62	.68
1.8	44	10	54	.68
1.9	34	8	42	.52
MIDDLE				
1.4	49	24	73	.50
2.1.iii	28	24	52	.08
3.3	35	23	58	.24
3.5	43	16	59	.54
HIGHER				
2.1.viii	33	28	61	.1
2.1.xi	31	27	58	.08
2.2	38	26	64	.24
2.3	45	14	59	.62
3.11	31	11	42	.4

DISCUSSION

A preliminary interpretation of these results is as follows:

There is a slight tendency for higher order skills to be more difficult. However, this tendency is not marked at all, thus, for example, lower order questions range from 42% to 89% in difficulty; higher order questions range from 42% to 64% in difficulty, and middle order questions range from 52% to 73% in difficulty. One of the lowest order questions, question 1.9, proves to be joint-most difficult with one of the highest order questions, number 3.11. In other words, counter-intuitively, there does not seem to be a progression in difficulty of items as the items progress up the scale of levels of comprehension and understanding from lower order through middle order to higher order. Item difficulty, in other words, does not seem to relate closely to the supposed level of comprehension, contrary to expectations.

If one inspects the discrimination indices for each item it is clear that, on the whole, better discrimination is achieved by lower order items than higher order items. Admittedly, the sample is small – five items in each – and there certainly is a range of discrimination indices, but if we ignore item 1.2, which it would be reasonable to do (given its high facility value one would expect relatively little discrimination to be possible anyway), then it is clear that high discrimination exists with lower order

questions and, on the whole, lower discrimination with higher order questions.

Looking at the number of students in the top and bottom groups who actually respond correctly to each item, there does appear to be a tendency for the differences between top and bottom group to be levelled out by higher order questions. In other words, the poorer discrimination of higher order items is influenced both by the bottom group doing somewhat better on some items than they did on lower order items, and by the top group doing somewhat worse on higher order items than they did on lower order items. Furthermore, what seems to be happening is that weaker students overall do somewhat better on higher order questions than on lower order questions.

This finding (if indeed it is a finding, of course, and these data are somewhat limited) is somewhat surprising and needs some explaining. What follows is entirely speculative, but hopefully interesting. It could be that weaker students overall are less disadvantaged by higher order questions than by lower order questions. That is to say, perhaps lower order questions measure language ability whereas higher order questions might be said to measure something like cognitive skills, logic, reasoning ability and so on.

There is no reason to believe, to our knowledge, that those students who are linguistically weaker in this sample are necessarily weaker cognitively, that is that they would necessarily perform worse at higher order skills, say, in their native tongue. Indeed, if the two groups of students are roughly equal in cognitive skills (and there is no reason to believe otherwise) then perhaps higher order questions allow this equality to show, whereas lower order questions are more language-bound and, therefore, in this sense less "fair".

If we want to measure language ability rather than cognitive skills then, perhaps it would be better to design a test with lower order questions which will maximise the language content validity of the test, and minimise contamination from cognition or higher order skills.

Since the groups are divided based upon their performance upon the test as a whole, the lower order questions must correspond more to what the test as a whole is measuring than higher order questions. Hence the high discrimination. If the test as a whole measures lower order skills, then this would not be surprising. If, however, the test as a whole measures language and communication skills, which it is intended to do, then this is indeed an interesting finding and suggests that what was said above might be true.

Whilst emphasising that this is entirely speculative, there are possible consequences that make sense. Depending upon the purpose of the test, it might perhaps be possible to get a better estimate of a student's reading abilities, rather than language

abilities from higher order questions than from lower order questions. Similarly, a better estimate of a student's cognitive abilities might be achieved on higher order questions in a foreign language than on lower order questions in a foreign language, because lower order questions might appear to be more language bound. However, if one's interest is language – testing language ability – it might be more appropriate to use lower order questions. But then the question arises – why is one interested in testing language for language's sake? Language is surely a means to an end, something to be used in order to understand texts. Such an understanding must surely importantly include understanding the main idea of the text, interpreting the text, relating the text to its context and the outside world, making judgements about the text, evaluating and synthesising the text and so on. In other words, language is a means to the end of communicating at high levels of cognition.

Finally, one should not perhaps believe that students with lower language levels are incapable of answering higher order questions. In other words, one should not be inferring from poor performance on lower order questions an inability to perform well on higher order questions. This would suggest that there is no implicational scale in reading in a second language such that one needs lower order abilities before one can progress to higher order questions.

REFERENCES

- Adams-Smith D.E. (1981) Levels of Questioning: Teaching Creative Thinking in ESP. *English Teaching Forum* January 1981, 15-17, 21.
- Bloom B.S. (1956) *Taxonomy of Educational Objectives. Book 1 Cognitive Domain*. London: Longman Group Limited
- Branford J.D., B.S. Stein and T. Shelton (1984) Learning from the perspective of the comprehender. In Alderson, J. Charles and A.H. Urquhart (Eds) *Reading in a Foreign Language* London: Longman, 28-44
- Carroll J.B. (1969) From comprehension to inference. In M.P. Douglas (Ed) *Thirty-Third Yearbook, Claremont Reading Conference* Claremont, Calif: Claremont Graduate School, 39-44
- Carroll J.B. (1971) Defining language comprehension: Some speculations. Research Memorandum 71-9. Princeton, N.J.: Educational Testing Service
- Davis F.B. (1944) Fundamental factors of comprehension in reading. *Psychometrika* 9, 185-197
- Davis F.B. (1968) Research in comprehension in reading. *Reading Research Quarterly* 3, 499-545

- Gray W.S. (1960) The Major Aspects of Reading, In H. Robinson (Ed) *Sequential Development of Reading Abilities*, Supplementary Educational Monographs No. 90. Chicago: University of Chicago Press, 8-24
- Lunzer E. and K. Gardner (Eds) (1979) *The Effective Use of Reading*. Heinemann Educational Books
- Lunzer E., M. Waite and T. Dolan (1979) Comprehension and Comprehension Tests. In Lunzer and Gardner (1979), 37-71
- Munby J.L. (1978) *Communicative Syllabus Design*. Cambridge: Cambridge University Press
- Spearritt D. (1972) Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Reading Research Quarterly* 8, 92-11
- Thorndike R.L. (1971) Reading as reasoning. Address delivered to Division 15, American Psychological Association, Washington D C
- Thurstone L.L. (1946) Note on a reanalysis of Davis's reading tests. *Psychometrika* 11, 185-188