COGNITIVE ABILITIES: VERBAL COMPREHENSION AND ITS MARKER TESTS

Joseph P. Boyle

Abstract

In all classifications of cognitive abilities a Verbal Comprehension factor is prominent. The nature of this factor is complex, and attempts to analyse it (Thurstone, 1938; Carroll, 1941; Guilford, 1972) have made distinctions between receptive ability (comprehension) and productive ability (fluency), and between isolated words (vocabulary) and words in context (discourse). The validity of these distinctions is discussed. The second part of the article examines what type of tests are the best markers for the Verbal Comprehension factor. The Educational Testing Service's kits of factor referenced cognitive tests (1954, 1963, 1976) give only single-word vocabulary tests, despite suggestions by Cattell (1971) and Carroll (1974) that other types of tests should also be used, and despite increasing doubts, especially in some ESL/EFL circles, about the validity of discrete-point language testing. The final part of the article describes an experiment with Hong Kong Chinese students. The results support the use of single-word Vocabulary tests as reliable markers of the Verbal Comprehension factor, but also support the contention that future Kits of The Educational Testing Service should include verbal tests of a more varied nature.

Joseph Boyle has an M.A. (Oxford) in English Language and Literature, Dip. E.S.L. (Leeds) and a Ph.D. in the area of listening comprehension from the University of Hong Kong. He has taught in Europe, South America, and Asia (India, the Philippines, and Hong Kong). He is a lecturer in the Chinese University of Hong Kong.
Cognitive Abilities and the Analysis of Verbal Comprehension

The classification of human cognitive abilities has been a task which has engrossed philosophers and psychologists for centuries. In the first part of this century important advances were made by such as Spearman (1904), Thorndike (1921), Thurstone (1938), and Burt and Jones (1942). Spearman divided human abilities into a general factor ("g"), plus specific factors ("s"), describing the specific factors as Verbal, Numerical, Mechanical, Attention, and Imagination. On the same lines, Burt and Jones postulated a general factor, plus specifics, which they labelled Verbal Reasoning, Language Usage, Numerical Ability, Mechanical Reasoning, Abstract Reasoning, Space Relations, and Clerical Speed/Accuracy. Thorndike preferred to describe human abilities in terms of a few large group factors, including Verbal, Mathematical, Dexterity. Following Thorndike, Thurstone distinguished six Primary Mental Abilities: Verbal, Word Fluency, Numerical, Spatial, Memory, and Reasoning.

A glance at these lists reveals the presence of a Verbal factor in all of them. Cattell (1971) notes how the Verbal factor had a special status for many psychologists: For Spearman it was a "hierarchy breaker"; for Burt and Vernon "almost a general factor"; for Thurstone "an emphatic primary". In Cattell's own scheme for classifying human abilities, his Universal Index, the first ability on the list is Verbal Ability. Guilford (1967) too, in his Structure of Intellect model, makes much of the Verbal factor.

As psychological knowledge became more refined, the classification of cognitive abilities became more complex. There has been controversy among psychologists working in this area about which type of classification is appropriate for a scientific description of cognitive abilities, some suggesting a matrix type of model (Guilford, 1967), as in chemistry, others preferring a hierarchical type of model (Royce, 1973), as in biology. Guilford’s Structure of Intellect model, with
Cognitive Abilities

its ingenious and detailed complexity, challenged the simplicity of previous models, but was itself challenged by Cattell (1971:55) who claimed that it was unconvincing, contained "too many arbitrary features", and was based on a method of factor analysis which was inappropriate.

The debate between Guilford and Cattell grew more fierce during the 70's, and reached a point in the 80's where even the titles of their articles reflected their personal animosity: Guilford (1980) scorned Cattell's idea of two general factors of intelligence, Gf (Fluid intelligence) and Gc (crystallised intelligence), in an article entitled, "Fluid and crystallised intelligence: two fanciful concepts", Cattell, with his colleague Horn, (Horn and Cattell, 1982) replied with: "Whimsy and misunderstandings of Gf-Gc theory: a comment on Guilford". Others have been less virulent, and possibly more effective, in their criticism of both Guilford (Carroll, 1968) and of Cattell (Eysenck, 1972).

The Verbal Comprehension Factor

Whatever disagreement there has been about the classification of cognitive abilities, there has been no disagreement about the central position of the Verbal Comprehension factor. Ekstrom, French, Harman and Dermen (1976) say such a factor has been mentioned explicitly in at least 125 published studies. Northrop (1977), in a history of the Verbal Comprehension factor, says it is a factor which does not easily break up into sub-factors. Nevertheless, there have been several attempts to refine and further analyse the Verbal factor.

Thurstone (1938) in his list of Primary Mental Abilities, included a Verbal factor (V) and a Word Fluency factor (W). He described the V factor as being logical in character, dealing with the understanding of ideas in discourse, rather than of isolated words. His W factor was associated with single, isolated words. Carroll (1941), in a study of the Verbal
Comprehension factor, re-analysed Thurstone’s data and concluded that the V and W factors could each be further analysed. He divided the V factor into two, a C and J factor. Factor C was related to the understanding of vocabulary, the “verbal tokens” which underlie ideas, rather than, as Thurstone had said, the comprehension of ideas as they occur in discourse. This latter, in Carroll’s analysis, belonged to a J factor. Thurstone’s W factor was also divisible, according to Carroll, into an A and an E factor. Factor A was characterised by speed of word association, whereas factor E influenced the speed of production of coherent discourse.

Both Thurstone and Carroll, therefore, wished to make practical distinctions between receptive (comprehension) and productive (fluency) ability, and between words in isolation (vocabulary) and words in coherent text (discourse), Thurstone suggesting a simpler V/W distinction, and Carroll preferring the more refined analysis into factors C and J (=V) and factors A and E (=W).

Guilford (1972:132) also believed that Thurstone’s V and W needed further analysis, and “should be regarded as verbal composites, each a confounding factor that represents a number of semantic abilities”. Using the terms of his own Structure of Intellect model, he claimed that Thurstone’s V factor represented no less than seven distinguishable factors. Inherent in Guilford’s analysis too, were the distinctions between receptive and productive, and between isolated words and words in discourse. Both these distinctions, however, need examining.

The Receptive/Productive Distinction

One of the most widely accepted theories of the comprehension process is the analysis-by-synthesis model of Halle and Stevens (1964), which has survived, with modifications, for the past twenty years. According to this model, the listener generates a sentence on the basis of a hypothesis
about the grammatical structure of the incoming signal, and checks this against the actual input. Miller (1964: 30) describes the process: The first stage is a guess about the stream of incoming sound; a response to this is then generated; the first guess may well be wrong, so another guess is made which will probably be closer; the listener keeps trying till a satisfactory match is obtained.

Though the active, constructive nature of the analysis-by-synthesis (A-by-S) mode seems to fit our internal awareness of the comprehension process, it was challenged by Katz and Postal (1964: 167). Their objection was that in order to provide the syntactic analysis of even fairly short sentences, the number of independent operations required would be so enormously high that a human brain could not be expected to perform the analysis even in a lifetime. The trial-and-error process would have too many errors and could not keep up with the incoming signals.

Neisser (1966) attempted to confront this objection with two suggestions. His first suggestion was that the incoming flow of speech passes through a “filter” system which segments the flow, extracts a few distinctive features, and tentatively recognises some of the constituent elements or units. What comes through this filter system is the raw material for the listener’s construction or synthesis of an internal message to match the actual input. Neisser’s second suggestion was that the constructive process is not aimlessly trial-and-error, but bases its construction on contextual clues. The context thus ensures that the most probable “fit” will be tried first, and since this will often be the correct one, the trial-and-error process will be characterised more by success than by error.

Another objection raised against the A-by-S model was that of Straight (1976) who pointed to the ability people have to interpret input that they cannot themselves produce. He claimed that this argued strongly against the blurring of the distinction between comprehension and production.
He also rejected Neisser's (1966) and others' revised models on the grounds that they all necessitated some sort of structural analysis prior to the matching procedure, an analysis which itself required explanation.

This line of objection had really been met earlier by Stevens and Halle (1967) who pointed out that the criteria employed in the matching operation were probably not very stringent. Neisser (1976:36) agreed that the matching constructions, i.e. the guesses, would need to be fairly open and not too specific. Cooper (1979:40) describes the input data as "somewhat rough and noisy", and suggests that the matching process in comprehension is based on "rather crude information".

More recent attempts to analyse the comprehension process still favour active, constructive models akin to the A-by-S model, and accept that reception and production of meaning cannot be rigidly separated. Rivers (1980:2) describes listening comprehension as a process of selecting and matching our selection against the incoming signal. In an earlier work Rivers (1976:133-137) gives a detailed analysis of the comprehension process. She distinguishes three stages: the first, a forming of rough impressions; the second, a more detailed attempt to segment and recognise lexical and syntactic patterns; the third, a recording of the material to suit the requirements of long-term memory.

In Abbott's model (Abbott and Wingard 1981), short term memory is actively applied to turn the stream of incoming speech into internal meanings. It holds stretches of speech while it operates on them. Features of a present stretch enable predictions, often very accurate, about future stretches to be made. These predictions are a vital component of speedy comprehension. What Abbott terms "strategies for understanding", working on both the incoming data and the surrounding context, then turn the stretches of speech into meaningful phrases. These are passed on to the long term memory where they are attached to an already existing network of meaning.
Cognitive Abilities

It is apparent from these attempts to analyse the comprehension process that the distinction receptive/productive, while it may often be useful for practical purposes, is not really as clear-cut and straightforward as it may at first seem.

The Vocabulary/Discourse Distinction

The second distinction commonly made when analysing the Verbal factor is a distinction between words in isolation and words with surrounding context. Vocabulary knowledge might be defined roughly as “knowing the meaning of the words”. The “meaning of a word” may sound a simple enough notion, and dictionaries are based on the idea that words can be defined in terms of necessary properties. However, linguistic philosophers in the 50’s, like Ryle (1951) and Wittgenstein (1953), showed the difficulty of defining words in terms of necessary properties, offering as examples such everyday words as “games” and “work”. Osgood, Suci, and Tannenbaum (1957) also showed, by means of their semantic differential vocabulary test, the complex nature of the “meaning” of words, especially when the word has affective connotations, like the word “mother”.

Clark and Clark (1977:45) observe that although meaning obviously plays a central part in comprehension, it has been given scant attention in the psychology of language. They suggest, among the reasons for this, the intrinsic difficulty of the concept, and the lack of an agreed framework in which to consider meaning.

For some kinds of words a neat, brief definition, like a dictionary entry, works well enough. But the meaning of other words can be approached only by a lengthy discourse, more like the entry in an encyclopedia. Hence the distinction made in semantic studies between “componential analysis”, which describes the meaning of words more like a dictionary entry, and “procedural analysis”, which can be compared more with an entry in an encyclopedia (Moates and Schumacher, 1980).
The problem is that words are not self-sufficient, isolated entities. The “meaning of a word” includes the relations of that word with others. Semantic field theory is based on the fact that the vocabulary of a language does not consist of a random list of words, but of networks of relations between words (Channel, 1981). Wilkins (1972:124) notes that the traditional view of “the meaning of a word” is that each word “has” a meaning, which is the primary thing, and that any relations between the meanings of words ultimately derive from this basic meaning which the words have in themselves. Wilkins suggests that the situation is really the reverse, that words cannot be understood in isolation, and that it is precisely the complex relations between words which determine the semantic structure of the language.

It can be seen, therefore, that the distinction between vocabulary (isolated words) and discourse (words in context), though necessary for practical purposes, like the receptive/productive distinction, is not as simple and clear-cut as it may at first appear to be.

Part 2:
Marker Tests for the Verbal Comprehension Factor

One of the reasons why Carroll, Guilford, and others felt obliged to subdivide Thurstone’s V and W factors was that the marker or reference tests which loaded on the factors were of very diverse types. On his V factor, for example, all of the following tests loaded: Vocabulary, Grammar, Spelling, Inventive Synonyms (give two words the same in meaning as the test word), Inventive Opposites (give two words opposite in meaning to the test word), Reading/Proverbs (select from alternative sentences the one which means the same as a given proverb), Reading/Quotations (select from alternative sentences the one which means the same as a given quotation).

A point of particular interest in Thurstone’s data was
the type of test which loaded most heavily on the Verbal factor, in other words, the type of test which would be the best marker or reference test for the factor. The two best candidates seemed to be Vocabulary test or tests of Reading Comprehension. In Thurstone’s original analysis the two Vocabulary tests had loaded on the Verbal Comprehension factor at .38 and .40—substantial, but not notably high loadings. However, using a different method of rotation on Thurstone’s data, others found vocabulary to be much more prominent than in the original analysis. Zimmerman (1953) found the loadings for the two Vocabulary tests on the Verbal Comprehension factor increased to .68 and .76, while Wrigley, Saunders and Newhaus (1958) found the loadings for the Vocabulary tests .74 and .93—all very high loadings.

Northrop (1977), reviewing a large number of factor analytic studies which had found a Verbal Comprehension factor, concluded that the purest measures of the factor were Vocabulary test. Reading Comprehension test, on the other hand, seemed to sample broader aspects of verbal ability, like ability to extract the main idea, or to make an inference. Northrop (1977:7) gave a list of several studies in which the loadings of the Vocabulary tests on the Verbal Comprehension factor were higher than the loadings of the Reading Comprehension tests on the same factor. Four examples from the list illustrate the difference:

<table>
<thead>
<tr>
<th>Study</th>
<th>Vocabulary</th>
<th>Reading Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fruchter (1952)*</td>
<td>.71</td>
<td>.56</td>
</tr>
<tr>
<td>French (1957)</td>
<td>.66</td>
<td>.41</td>
</tr>
<tr>
<td>Kelley (1964)</td>
<td>.60</td>
<td>.48</td>
</tr>
<tr>
<td>Very (1967)</td>
<td>.89</td>
<td>.76</td>
</tr>
</tbody>
</table>

*all citations in Northrop, 1977
It is interesting that all three editions of the Educational Testing Service’s (ETC) kits of factor referenced cognitive tests (Ekstrom, French, Harman and Dermen, 1976; French, 1954; French, Ekstrom and Price, 1963) give only vocabulary tests of the single-word, multiple-choice-type as reference tests for the Verbal Comprehension factor. This is somewhat surprising since the factor is defined in the kits broadly as, “the ability to understand the English language”. Carroll (1974), in a preparatory paper for the third edition of the ETS Kit, commented on this and suggested that a more diversified set of tests for the Verbal Comprehension factor might be more appropriate. Cattell (1971) too did not think Vocabulary tests alone were adequate to mark the Verbal Comprehension factor, and he offers a fairly wide set of tests to measure the ability, including Vocabulary, Reading Comprehension, and Syntax.

Experiment with Chinese Students

In order to obtain further independent evidence on the question of what type of tests could best act as markers of the Verbal Comprehension factor, a study was conducted in Hong Kong. The subjects were 285 students from the Chinese University, 144 males and 141 females, aged 18 to 20. All spoke Cantonese as their mother tongue, and had studied English in school for ten or more years.

The investigation of the best marker tests for the Verbal Comprehension factor was part of a broader experiment, involving a large battery of tests, and using the method of factor analysis. It is beyond the scope of this article to discuss all the different tests and the reasons for their inclusion in the analysis. A brief description, however, of the eight verbal tests which are immediately relevant to the present discussion may be helpful.

1. Vocabulary A. This was taken from the English Language Battery (ELBA) (Ingram, 1984). It is a standard
single-word, multiple-choice vocabulary test, with a test word and four alternative choices.

2. **Vocabulary B.** In this test, adapted from the *Sequential Tests of Educational Progress* (Educational Testing Service, 1979), another single-word Vocabulary test, the test word was in English, but the multiple-choice alternatives were translated into Chinese.

3. **Vocabulary C.** This was the *English Picture Vocabulary Test* (Brimer and Dunn, 1968). One word is given in English, and the candidate must choose which of four pictures best represents the word.

4. **Reading Comprehension A.** This was a standard TOEFL Reading Comprehension test with global comprehension questions on several paragraph-length passages.

5. **Reading Comprehension B.** This tested more detailed comprehension, the candidate having to choose from four sentence-length alternative answers.

6. **Dictation A.** This was a narrative passage, adapted from a newspaper article, about a sailing trip in a Chinese junk from Singapore to Java. The passage was read right through, then repeated in sections of about ten words at a time, then read right through again.

7. **Dictation B.** This was like Dictation A in length and manner of testing, only the style was very different, a literary description of a character in a novel.

8. **Cloze.** This was a standard cloze with deletions about every ninth word, the deletions being made rationally, not randomly, and scoring done on the acceptable alternatives method, rather than on the exact word method.

**Results**

As is usual in factor analytic studies involving verbal tests, a first large factor emerged, which could be considered as
the Verbal Comprehension factor. The loadings on this factor of the eight tests are given below:

Table 2
Loading of Tests on the Verbal Comprehension Factor

<table>
<thead>
<tr>
<th>Test</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vocabulary A</td>
<td>.53</td>
</tr>
<tr>
<td>2. Vocabulary B</td>
<td>.58</td>
</tr>
<tr>
<td>3. Vocabulary C</td>
<td>.65</td>
</tr>
<tr>
<td>4. Reading Comp A</td>
<td>.53</td>
</tr>
<tr>
<td>5. Reading Comp B</td>
<td>.45</td>
</tr>
<tr>
<td>6. Dictation A</td>
<td>.69</td>
</tr>
<tr>
<td>7. Dictation B</td>
<td>.73</td>
</tr>
<tr>
<td>8. Cloze</td>
<td>.63</td>
</tr>
</tbody>
</table>

Factor loadings of .4 and above are normally considered of great interest.

It can immediately be seen that all the single word vocabulary tests loaded substantially on the Verbal Comprehension factor, and therefore can be said to be good marker tests of the factor. They seem to be better markers than the reading comprehension tests, whose loadings, though fair, are not so high. However, the Cloze has a higher loading than two out of the three vocabulary test, while the highest loadings are achieved by the dictation tests.

The conclusions therefore of this study are as follows: First, single-word vocabulary tests can still be considered good markers of the Verbal Comprehension factor. Second, however, the best marker tests appear to be tests such as the dictation tests which call on a wider range of verbal abilities. This second conclusion is in agreement with Oller and Perkins (1980) and others who prefer integrative tests to discrete-point test. The first conclusion – the abiding efficacy of single-word Vocabulary tests – might constitute a call to caution for language teachers lest, in overemphasising the communicative and discourse aspects of language, they reject too readily a type of test which is considered by psychologists...
Cognitive Abilities

to be extremely reliable.

On the other hand, the message which emerges for psychologists investigating the Verbal factor in cognitive abilities, is to diversify the types of test they use, instead of relying solely on single-word multiple-choice vocabulary tests as they have tended to do in the past. While this type of test is attractive in its simplicity and speed of administration, and while it has proven reliable as a marker for the Verbal Comprehension factor, nevertheless it is hardly adequate as the only marker of a factor broadly defined as "the ability to understand the English language".

Conclusion

Within the wider framework of attempts to classify human cognitive abilities, the Verbal Comprehension factor has been discussed. Attempts to analyse this factor have been described, and some distinctions commonly made have been seen to be inadequate.

Given the complexity of the factor, it is hardly surprising that there is disagreement on what kind of tests are the most appropriate marker tests for Verbal Comprehension. The practice among some psychologists of using only single-word Vocabulary tests would be challenged by many language teachers, who themselves, however, should not be too cavalier in dismissing this type of test from their test batteries.

References


Cognitive Abilities


203
The RELC Journal: A Journal of Language Teaching and Research in Southeast Asia is one of the professional publications of the Southeast Asian Ministers of Education Organization (SEAMEO) Regional Language Centre (RELC) which is located in Singapore. The main purpose of the journal is to disseminate information and ideas on theories, research, methods and materials related to language learning and teaching. Articles covering a wide range of topics in applied linguistics appear in its pages semiannually in June and December. Special features include information on current research projects in Southeast Asia, book reviews and review articles on topics related to language teaching. In its thirteen years of publication, the journal has reached a wide audience in eighty countries and has been rated as one of the leading journals in applied linguistics throughout the world. Scholars of recognized stature have used it as a means of presenting their findings and obtaining feedback on important issues in language and linguistics.

GUIDELINES

A companion publication, Guidelines, provides language teachers with practical and innovative techniques to enliven and enrich teaching in the language classroom. Each issue focuses on a particular theme.

The Journal and Guidelines are published twice a year and a subscription may begin with the June or December issue. The annual subscription is US$9.00 for each title. Individual copies may be purchased at US$6.00.

Please send your order to The Publications Officer
SEAMEO Regional Language Centre
RELC Building
30 Orange Grove Road
Singapore 1025
REPUBLIC OF SINGAPORE
The highly motivating EFL/ESL program for teaching children ages 3-6 communication skills in English

My English Book Series offers:

- Complete, sequential program with a teaching approach to ensure success
- Clear organization to simplify teaching and testing
- Four-step learning sequence to increase student participation

For more information, contact:
Haruhiko Kuwashima • 203 Royal Maison Tsukushino
4-22-4 Ogawa, Machida-shi • Tokyo 194, Japan

Scott, Foresman International Division
1900 East Lake Avenue, Glenview, IL U.S.A.
Telephone (312) 729-3000 • Telex: 724371
Cable: SCORESMAN, Glenview, IL U.S.A.