

Construct Validity of L2 Reading Comprehension Skills

Hossein Farhady

Iran University of Science and Technology

Gholam Reza Hessamy

University of Tehran

Abstract

The research on the nature of the underlying reading sub-skills is still unsettled. Nor is there any consensus over the divisibility of reading into a set of component skills. Improving construct validity of tests of L2 reading comprehension requires further investigation on the variables underlying reading ability and the possibility of operationalizing the underlying traits. To this end, a Test of Reading Sub-skills (TRS) was developed which operationalized the components of the Comprehensive Taxonomy of Reading Sub-skills (CTRS) derived from major taxonomies in the literature through a detailed comparative analysis of their proposed components. The TRS employed a wide variety of test methods in two general categories of the Expected Response Format, namely, 'Selected Response' and 'Constructed Response'. Data was collected on the performance of 1606 Iranian EFL learners including English majors and non-English majors, each subject randomly answering one of the eight forms of the TRS. Data were analyzed using Exploratory and Confirmatory Factor Analysis. The results showed that L2 reading ability is composed of a number of underlying latent traits or macro-skills. It is maintained that a componential approach to reading assessment and pedagogy should be informed by empirical evidence on the nature of variables underlying reading ability.

Key Word: Construct Validity, Reading Comprehension, Iran EFL Learners.

Introduction

In the context of foreign or second language learning, reading is probably the most frequently used channel of communication in the target language and the most urgent need for learning the language. This is more evident in the case of English as a foreign language, i.e., learning English in a context in which English is not normally spoken in the society, owing to the fact that English is the unrivaled international medium of written communication. According to Grabe and Stoller (2002, p. 2):

Reading in second language (L2) settings continues to take on increasing importance. The overwhelming majority of societies and countries around the world are multilingual, and educated citizens are expected to function well in more than one language. L2 reading ability, particularly with English as the L2, is already in great demand as English continues to spread, not only as a global language but also as the language of science, technology and advanced research. Many people in multilingual settings need to read in an L2 at reasonably high levels of proficiency to achieve personal, occupational and professional goals.

Generally speaking, research on the nature of reading has focused on two main perspectives, namely, process-oriented and product-oriented (Urquhart and Weir, 1998). While the process-oriented view explores what actually happens in the mind of the reader during reading, the product-oriented perspective examines the performance of the reader to shed light on the nature of the underlying latent abilities.

An important part of reading research and theory is the study of the nature of underlying variables that together make up fluent readers' ability to understand different levels of meaning in a text. This approach to the study of reading is based on a view of reading as separable and distinguishable underlying components (Hoover and Tunmer, 1993). The

rationale behind the study of the underlying skills is mainly pedagogical. It is hoped that by breaking down the reading process into its components, they can be taught systematically and tested separately with a further assumption that the composite result would represent the overall reading ability of the learner (Weir and Porter, 1994; Urquhart and Weir, 1988; Weir et al. 2000).

Notwithstanding the fact that a componential approach to describing reading ability has valuable applications to the practice of teaching and testing, research on the divisibility of the construct of reading into separately identifiable components is inconclusive. In this regard, Weir and Porter (1994) provide a good review of the quantitative and qualitative studies on the divisibility of reading. As Lumley (1993, p. 211) states, "the very concept of describable reading subskills is highly controversial, despite its frequent occurrence in the areas of both syllabus design/materials preparation and test construction". Alderson (2000, p.10) argues along the same line by saying that "there is a considerable degree of controversy in the theory of reading over whether it is possible to identify and label separate skills of reading".

As expressed by Urquhart and Weir (1998, p. 47), "provided that it can account for the observed data, a model with fewer variables is preferable to one with more". Weir et al. (2000) believe that reading is at least a "bi-divisible process" in terms of skills and strategies. However, this is only one way to divide reading into components. In fact, a number of two-component models have been suggested so far, among them Hoover and Tunmer's (1993) "simple view of reading" is probably the most advocated one in recent years (Grabe, 1997).

Three-component models are best known by those of Coady (1979) and Bernhardt (1991). Coady's model includes conceptual abilities, process strategies and background knowledge. Bernhardt's model is

composed of language, literacy and world knowledge. There are still studies which found more underlying variables based on empirical data (e.g. Moeini Asl, 2002; Daftari Fard, 2002). Carroll (1993) reanalyzed 30 factor analytic studies and found four factors underlying reading ability: (general) reading comprehension, special reading comprehension, reading decoding and reading speed. Buck et al. (1997) also refer to studies using factor analysis and point to the debate on the interpretation of analyses.

There are also studies that have investigated the components of reading ability with special attention to underlying linguistic skills as predictors of reading comprehension test performance. Shiotsu (2004) investigated the relative contribution of grammatical and lexical knowledge to reading comprehension ability of L2 readers and found a greater role for syntactic knowledge than vocabulary in predicting reading test performance. Alavi (2004) also found grammatical ability as measured in the 'Written Expression' subtest of the TOEFL as a better predictor of reading test performance than vocabulary knowledge.

Research on divisibility of the construct of reading is yet to be pursued more rigorously. Multi-component taxonomies of reading skills have been recommended and used for teaching and testing purposes for a rather long time. Well-known international proficiency tests claim to use lists of reading skills as a guiding framework for designing test items of different nature (e.g. IELTS and TEEP). However, there are some major criticisms against these taxonomies. Williams and Moran (1989), cited in Urquhart and Weir (1998), refer to the lack of consensus regarding the number and kind of sub-skills to be included and the terminology used for the components. It is common to find components which actually refer to the same ability but differently labeled, and this would naturally be a source of confusion. Furthermore, as pointed out by Urquhart and Weir (1998), some taxonomies include very inclusive categories, e.g. Grabe

(1991) mostly refers to general knowledge areas rather than specific skills. Alderson (2000) also mentions the lack of clear definitions of the components and maintains that they are rarely as discrete as the taxonomies would suggest. More specifically, he argues that it is almost impossible to isolate what skills are operationalized by what test items, and that analysis of test performance does not support such a separation of skills. Alderson believes that reading involves several overlapping 'skills' which are used in connection with each other as necessary. Alderson (2004) questions the plausibility of characterizing individual test items as measuring particular reading sub-skills separately. He (*ibid*: 15) maintains that "generalizations about what skills reading test items might be testing are fatally flawed". This can further raise the question of construct validity of measures of reading ability.

As noted by Douglas (1995, p.169), "In language testing, construct definition is clearly important because it is only by reference to underlying constructs that test performance can be interpreted". Language testers concerned with the validity of their tests of reading cannot and should not go along with unsubstantiated assumptions about the nature of the construct. Tests of reading are valid to the extent that their underlying theory is validated.

This study was designed to find an answer to the question regarding the nature and the componentiality of L2 reading comprehension ability.

Method**Participants**

The participants in this study were 1606 Iranian EFL learners including junior and senior English majors (EM) and non-English majors (NEM). About 70% of the subjects were EMs and 76% of the participants were females. The subjects were between 15 and 36 years of age. The sample of EMs included students of TEFL (18%), Translation (17%) and literature (35.5%) in different State and Azad Universities in Tehran, Mashhad, Kermanshah, Yazd, Torbat-e-Heydariyeh and Quchan. The NEM group was composed of students taking IELTS, TOEFL, FCE and CAE preparation courses and other general English courses at advanced levels in a number of private institutes.

Instrumentation

The main instrument of this study was a test of reading sub-skills (TRS). The TRS operationalized 28 sub-skills of reading derived from a detailed comparative analysis of components proposed in the major taxonomies in the literature. The theoretical definitions extracted from the literature were provided for each sub-skill as a guiding framework for operationalizing the components. Table 1 presents the components of the CTRS used as the theoretical model for the TRS.

Table 1
The Components of the CTRS

1. Deducing word meaning through word formation
2. Deducing word meaning through contextual clues
3. Understanding syntactic relations
4. Understanding conceptual meaning
5. Understanding grammatical cohesion
6. Understanding lexical cohesion
7. Understanding rhetorical organization
8. Understanding functional value
9. Understanding explicitly stated information
10. Understanding information clearly stated but in paraphrase
11. Understanding main idea and supporting details
12. The ability to make propositional informational inferences
13. The ability to make propositional explanatory inferences
14. Understanding figurative language
15. Understanding presuppositions underlying the text
16. Inferring what preceded
17. Transcoding diagrammatic display to writing
18. Transcoding writing to diagrammatic display
19. Evaluating inferences
20. Evaluating generalizations
21. Recognizing textual inconsistencies
22. Summarizing
23. Selective extraction of relevant points
24. Restating or Paraphrasing
25. Drawing conclusions
26. Predicting what follows
27. Choosing an appropriate title
28. Process analysis

The test operationalized each sub-skill in two general categories of the expected response format, namely 'Selected Response' (SR) and 'Constructed Response' (CR) using a large variety of test methods including multiple-choice, multiple-matching, dichotomous, multiple-choice rational cloze, rational cloze, short answer question, sentence/table/diagram/summary completion, and editing. The test was reviewed by experts, trialed twice with subjects similar to the target population, and revised for a number of times before it was used for the final administration. The final form of the test consisted of 184 items arranged in 8 booklets each to be answered in 60 minutes and distributed randomly among the subjects (some examples of the items with their specifications are provided in the Appendix).

The second instrument was a sample academic reading module of the IELTS which was used as a criterion measure administered to a portion of the subjects ($n=321$) in order to get an index of empirical validity of the TRS.

Procedure

The TRS was administered in the regular class time of the universities and institutes that cooperated in this study. Due to certain limitations in data collection, no subject answered more than one booklet of the TRS. The booklets were administered randomly among the subjects in each class. The criterion measure, however, was given to a portion of the subjects who were both EMs and NEMs to be answered in a separate time setting.

Data Analysis

The data were analyzed through the following steps. First, descriptive statistics was computed for all the measures of the study. Correlation analysis was conducted between the eight forms of the newly developed test and the criterion measure to check the empirical validity of the TRS. Then, to test the hypothesis set for the study, data were analyzed using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) to investigate the validity of the operationalized sub-skills and the latent structure of the test.

Results

Table 2 displays the descriptive statistics for the measures used in this study. As the table shows, the subjects did better on the IELTS than the TRS. In fact, the test run on the paired mean differences of the performance of the subjects who took the criterion measure (n=321) revealed significant differences in their performance on the IELTS and different forms of the TRS. This was probably due to different reasons including the fact that the TRS was actually more demanding as it utilized a wider range of reading skills and test methods and the majority of the subjects were more or less familiar with the format of the IELTS. The reliabilities of the different forms of the TRS appear to be moderate in all cases. It should also be noted that the reliabilities were affected by the homogeneity of the subjects which decreased the variance of the scores and accordingly the reliability estimates.

Table 2
Descriptive Statistics for the Measures of the Study

Tests	N	Max. Possible Score	Mean	Std. Deviation	Reliability (Alpha)
Booklet 1	198	23	7.05	3.70	.61
IELTS	43	38	21.72	4.03	.69
Booklet 2	232	27	10.46	3.80	.65
IELTS	57	38	26.47	4.52	.69
Booklet 3	215	21	6.87	2.82	.50
IELTS	38	38	21.55	4.37	.69
Booklet 4	203	20	8.27	3.68	.55
IELTS	34	38	25.94	5.59	.69
Booklet 5	206	23	8.04	3.14	.70
IELTS	42	38	28.83	5.04	.69
Booklet 6	172	22	7.80	3.35	.68
IELTS	25	38	23.96	5.60	.69
Booklet 7	189	25	7.34	3.27	.48
IELTS	31	38	19.96	5.49	.69
Booklet 8	191	23	8.51	3.61	.66
IELTS	39	38	24.53	4.44	.69

The correlation coefficients between different forms of the TRS and the IELTS academic reading were computed as presented in Table 3. The eight booklets of the TRS appeared to show rather high indices of empirical validity indicating that there is a good degree of common variance between each form of the TRS and the criterion measure.

Table 3
Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	IELTS & BOOK1	43	.79	.000
Pair 2	IELTS & BOOK2	58	.88	.000
Pair 3	IELTS & BOOK3	40	.80	.000
Pair 4	IELTS & BOOK4	35	.86	.000
Pair 5	IELTS & BOOK5	43	.86	.000
Pair 6	IELTS & BOOK6	32	.79	.000
Pair 7	IELTS & BOOK7	31	.83	.000
Pair 8	IELTS & BOOK8	38	.59	.000

In order to examine the psychological reality of the hypothesized sub-skills, the data were first analyzed by EFA to see how many underlying factors would emerge. Through EFA, 65 factors which were more than the number of hypothesized variables were extracted*.

However, a close examination of the factor loadings revealed the clustering of the variables that in the majority of the cases indicated meaningful relationships with regard to their underlying linguistic and processing characteristics. The following hypothetical latent abilities underlying reading ability were therefore identified: (1) *Inferential and interpretive skills*; (2) *Linguistic and textual contributory skills*; (3) *Understanding explicit information*; (4) *Process analysis*; (5) *Transcoding skills (diagram to writing or vice versa)*; (6) *Summarizing*; (7) *Selective extraction of relevant points*; (8) *Recognizing textual inconsistencies*.

Among the above latent abilities or macro-components, the 'inferential and interpretive skills' turned out to be the most frequently occurred combination of variables. The inferential skills required the reader to move from explicitly stated information or ideas to an inference or interpretation about the unstated information or ideas. Under this broad category which appeared in 17 factors, the sub-skills numbered 7, 8, 11, 12, 13, 15, 16, 25, 26, 27, 19, 20, 1, 2, and 14 appeared.

The second most frequently occurring cluster of variables belonged to those sub-skills which collectively could be labeled as the 'linguistic and textual contributory skills'. Four factors related to syntactic, conceptual and cohesive relations within the text appeared here. In one more factor, the linguistic and textual macro-component overlapped with the 'selective extraction of the relevant points'. By examining the items, it appeared to be the result of the sameness of the text to which the corresponding items were all related. By their very nature, the linguistic and textual skills underlie comprehension of the text at all levels of meaning. Under this category the sub-skills 3, 4, 5, and 6 were recognized.

It seemed quite reasonable to expect that the variables which required understanding and processing explicitly stated information function separately; however, the factor structure of the TRS did not support such an assumption. The main reason was probably related to the influence of different tasks on the clustering of variables. The effect of the similarity of the tasks could somehow indicate a different kind of processing but this conclusion could not be maintained without evidence on the processes that actually take place in the minds of the readers while doing different tasks. The sub-skills included here were 'summarizing', 'process analysis', 'selective extraction of relevant information', 'transcoding', and 'recognizing textual inconsistencies'. There were also two factors in which 'understanding explicit information and summarizing' and 'understanding explicit information and paraphrasing' made separate clusters. These were quite understandable as summarizing and paraphrasing are both concerned with explicit information. However, while paraphrasing did not emerge as a separate factor, summarizing items did appear in two factors separately. Consequently, based on the results of EFA, these macro-skills were also identified: (3) Understanding explicit information: sub-skills 9, 10, and 24; (4) Process analysis; (5) Transcoding: sub-skills 17 and 18; (6) Summarizing (7) Selective extraction of relevant points; (8) Recognizing textual inconsistencies.

While the above categorization of the underlying reading macro-skills cannot be found in its entirety in any single study, the findings of previous studies (e.g. Davis, 1968; Lumley, 1993; Weir et al. 2000; Moeini Asl, 2002; Daftari Fard, 2002) have more or less indicated the existence of most of these underlying skills but with slightly different characterizations. The macro-skills which appear in almost every empirical study are inferential skills and understanding explicitly stated information or ideas.

Since each of these macro-skills appeared in more than one factor, and inasmuch as the number of factors extracted was more than the number of the sub-skills in the theoretical model of the study, the data were analyzed using CFA with a 28-factor solution corresponding to the number of the components of the CTRS as the theoretical model underlying this study.

The results of the CFA to a large extent confirmed the previously hypothesized macro-skills while in a number of cases unexpected combinations of the sub-skills under a single factor caused some confusion. For example, seven factors clearly showed the clustering of sub-skills which were already designated as the inferential and interpretive macro-skill. Linguistic and textual contributory skills appeared exclusively in one factor while they clustered with the inferential skills in four other factors. This should not be surprising since as already mentioned linguistic skills are by nature underlying all text comprehension and contribute to understanding different levels of meaning in texts. Therefore, it seemed justifiable that linguistic and textual skills might function in close relationship to other skills. They were also clustered with the 'selective extraction of relevant points' in one factor as it happened in the EFA and already explained. Among other sub-skills, 'transcoding', 'selective extraction', and 'recognizing textual inconsistencies' each appeared in one separate factor which conformed to the previous hypothetical macro-components. 'Understanding explicit information' and 'paraphrasing' also appeared in one factor as it happened in the EFA.

However, some combinations of the sub-skills which dealt with explicit information in the text with those that required making inferences (factors 2, 6, 11 & 13) seemed quite confusing. There were other cases (factors 15, 16, 21, 24 & 28) which were hardly interpretable as they incorporated sub-skills belonging to more than two categories of the

hypothesized macro-components seemingly unrelated to each other. Nevertheless, there were positive indications of the similarity of the sub-skills which were already discussed as sharing the feature of understanding and analyzing explicit information albeit through utilizing different tasks. In eight factors different combinations of the sub-skills 9, 10, 22, 23, 24, 17, 18 & 28 could be found, although at times due to existence of other variables in the clusters, making logical interpretations seemed difficult. The evidence could somehow indicate that in spite of differences in the actual tasks, these sub-skills had some underlying commonality which was the centrality of analyzing explicitly stated information in the text as opposed to making inferences. Finally, contrary to the findings of EFA, the sub-skill of 'summarizing' did not emerge as a separate factor. Therefore, it seemed that summarizing had to be left out from the categorization of the macro-skills. Consequently, based on the results of the EFA and CFA, the following hypothetical macro-skills with the same sub-categorizations as proposed before can be suggested: (1) *Inferential and interpretive skills*; (2) *Linguistic and textual contributory skills*; (3) *Understanding explicit information*; (4) *Process analysis*; (5) *Transcoding*; (6) *Selective extraction of relevant points*; and (7) *Recognizing textual inconsistencies*.

Discussion

The central question in the product-oriented perspective to reading is the divisibility of the construct and the nature of the underlying abilities that together make up a fluent reader's ability to arrive at different levels of meaning in a text. The study of the nature of the product is particularly important for the test development process as tests of reading are concerned with the product of reading comprehension. However, the question of componentiality still needs further evidence to be settled. There is no consensus to date over the number and nature of underlying reading skills. While in the last thirty years or so there has been a

proliferation of theoretical taxonomies of reading skills and strategies with generous itemization of components in rather long lists, there has been a respectable body of empirical investigations that found one or at most two latent traits of reading ability (e.g. Rost, 1993; Lunzer et al., 1979). It is suggested that the sub-skills may become intermingled and undifferentiated as readers become skillful (Rost, 1993). Nevertheless, this study provided evidence on the performance of fairly advanced EFL readers of English which according to this explanation should reveal undifferentiated ability.

On the other hand, as long as the nature of tasks and test methods can affect subjects' performance and influence the measurement of particular sub-skills, the TRS utilized a large variety of test methods in order to provide a context in which the test was not biased due to the application of a particular test method. Nevertheless, the subjects may utilize unpredictable strategies or combinations of skills to arrive at an answer for which the involvement of another skill was envisaged by the test designer. This may account for unexpected loadings or functioning of some items in the analyses conducted. Furthermore, the difficulties inherent in specifying exactly what an item is testing always exist as reading behavior is so complex that it may require simultaneous involvement of a number of skills at a time. While it was attempted that the intended sub-skill measured by any item was the most important ability without which comprehension of the intended meaning was impossible, some overlap could naturally be expected. And finally, the effect of task characteristics cannot be ruled out as similar or separate functioning of some items can be as much an effect of task similarities or differences as the result of trait characteristics.

Considering all the potential sources of difficulty in operationalizing a language construct especially in the receptive mode, the loadings of the

factors in the two rounds of factor analysis were illuminating in that they clearly showed that the factor structure of L2 reading ability is so complex that one can hardly maintain a unitary view of the construct. Moreover, the emergence of macro-components which incorporated a noticeable number of hypothesized sub-skills clearly indicates the existence of the underlying latent traits. The other macro-components were the result of both trait and task effect as theoretically they could have subsumed under more inclusive macro-components. For example, 'selective extraction of relevant points', 'process analysis' or 'transcoding' are all concerned with understanding and analyzing explicitly stated information in one way or another, but they differ in terms of task characteristics, i.e. what is done with the information. Therefore, it can be concluded that they are better interpreted as combinations of trait and task characteristics.

The phenomenon of vocabulary or lexical knowledge emerging as a separate latent trait which is observed in many studies already discussed did not happen here as lexical skills measured in this study were different from 'vocabulary knowledge' or 'memory for word meanings'. The CTRS was a model of L2 reading comprehension skills only. As already noted, for theoretical and practical reasons, knowledge components and reading strategies were not included in this investigation.

Conclusion

This study provided empirical evidence for the validity of a number of important underlying traits in L2 reading comprehension. The empirical findings of the study are significant in the sense that the instrument used was based on a theoretical model which reflected all major taxonomies proposed by recognized reading specialists. The operationalized components were resonating the current collective thinking about the underlying reading component skills. Furthermore, the variety of test

methods used and the use of both selected response and constructed response items provided a context in which the instrument was not biased due to application of a single method.

Because of the appealing nature of the taxonomies of reading ability and the possibility of their direct application to teaching, material development and testing practice, it seems a promising task of researchers to investigate the psychological reality of the proposed reading sub-skills so that the practitioners can draw on empirically validated models and make informed choices on what to include in reading classes, textbooks, and tests.

Received 15 December 2004

Accepted 20 June 2005

Archive of SID

Acknowledgments

The authors would like to express their deep gratitude to the Provincial Office of Social Affairs and Elections in Khorasan Razavi for their invaluable assistance in collecting a large body of data.

References

- Alavi, S. M. (2004). Contribution of grammar and vocabulary knowledge to reading comprehension tests. *Pazhuhesh e-Journal*, No.17, Summer issue, Internet Page.
- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- , (2004). The shape of things to come: Will it be the normal distribution? In M. Milanovich (ed.), *Studies in language testing*, (pp. 1-26). Cambridge: UCLES.
- Bernhardt, E. B. (1991). A psycholinguistic perspective on second language literacy. In J. H. Hulstijn and J. F. Matter (eds.), *Reading in two languages, AILA Review 8* (pp. 31- 44). Amsterdam.
- Buck, G., Tatsuoka, K., and Kostin, I. (1997). The sub-skills of reading: Rule-space analysis of a multiple-choice test of second language reading comprehension. *Language Learning*, Vol. 43 No.3, pp. 423-466.

- Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press.
- Coady, J. (1979). A psycholinguistic model of the ESL reader. In R. Mackay et al. (eds.), (pp. 5-12).
- Daftari Fard, P. (2002). Scalability and divisibility of the reading comprehension ability. Unpublished MA thesis, Iran University of Science and Technology, Tehran.
- Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly*, Vol. 3, pp. 499-545.
- Douglas, D. (1995). Developments in language testing. *Annual Review of Applied Linguistics*, Vol. 15, pp. 167-187.
- Grabe, W. (1991). Current development in second language reading research. *TESOL Quarterly*, Vol. 25, No. 3, pp. 375-406.
- , (1997). Reading research and its implications for reading assessment. Paper presented at the 19th Language Testing Research Colloquium.
- Grabe W. & Stoller P. (2002). *Teaching and researching reading*. New York: Longman.
- Hoover, W. A. and Tunmer, W. E. (1993). The components of reading. In J. B. Thompson, W. E. Tunmer, and T. Nicholson (eds.), *Reading acquisition processes*, (pp.1-19). Clevedon: Multilingual Matters Ltd.

- Lumley, T. (1993). The notion of subskills in reading comprehension tests: An EAP example. *Language Testing*, Vol. 10 No.3, pp. 211-234.
- Lunzer, E., Waite, M., and Dolen, T. (1979). Comprehension and comprehension tests. In E. Lunzer and K. Guthrie (eds.), *The effective use of reading*, (pp. 37-71). London: Heinemann Educational Books.
- Moeini Asl, H. R. (2002). Construct validation of reading comprehension tests. Unpublished MA thesis, University for Teacher Education, Tehran.
- Rost, D. H. (1993). Assessing the different components of reading comprehension: fact or fiction. *Language Testing*, Vol. 10, pp. 79-92.
- Shiotsu, T. (2004). Investigating the relative significance of vocabulary breadth and syntactic knowledge in the prediction of reading test performance. Paper presented at 26th *Language Testing Research colloquium*.
- Urquhart, A. H., and Weir, C. J. (1998). *Reading in a second language: process, product, and practice*. New York: Longman.
- Weir, C. J., Huizhong, Y., Yan, J. (2000). An empirical investigation of the componentiality of L2 reading in English for academic purposes. Cambridge: UCLES.
- Weir, C. J. & Porter, D. (1994). The multi-divisible or unitary nature of reading: The language tester between Scylla and Charybdis. *Reading in a Foreign Language*, Vol.10, pp. 1-19.

Williams, E. and Moran, C. (1989). Reading in a foreign language at intermediate and advanced levels with particular reference to English. *Language Teaching*, Vol. 22, No. 4, pp. 217-28.

Appendix

Some examples of the items of the TRS (multiple-choice items)

Example 1: Sub-skill 1

The word 'anachronism' in paragraph 4 is closest in meaning to

- A. Atomic missile production
- B. Something placed in the wrong period of time
- C. Historical mistrust
- D. Political competition

Example 2: Sub-skill 11

What is the main idea of the third paragraph?

- A. Fingerprinting is not reliable anymore.
- B. Fingerprinting should be reexamined.
- C. Fingerprinting is not suitable for complicated cases.

- D. Fingerprinting should be substituted by some new techniques.

Example 3: Sub-skill 13

How does the writer reflect on the use of Modernist approach?

- A. With sympathy
- B. With regret
- C. With enthusiasm
- D. With concern

Example 4: Sub-skill 8

In the last sentence of paragraph 9 the writer tries to

- A. Speculate on something
- B. Predict something
- C. Give an advice
- D. Hypothesize a situation

Example 5: Sub-skill 15

Which of the following can be taken as an underlying presupposition of the passage?

- A. Our civilization is decadent.
- B. English language is deteriorating.

- C. Language is independent of individual writers.
- D. Language and politics are interrelated

Example 6: Sub-skill 16

Which of the following can best be the first sentence of the last paragraph of the passage?

- A. The fault does not lie with the schools alone, however.
- B. The universities are more responsible for not giving information about the job conditions.
- C. Nevertheless, the students themselves could work harder to find better jobs.
- D. The parents seem to be unjustifiably indifferent, too.

Example 7: Sub-skill 20

Which of the following generalizations can be arrived at based on the views presented in the last paragraph of the passage?

- A. Punishment should be morally justified.
- B. Public execution can suppress violent tendencies in the society.
- C. Physical punishment should be abolished.
- D. Punishment should be deterrent of further crime

Example 8: Sub-skill 25

It can be concluded that the statistics on the coincidence of prints are

- A. not reliable in practice.
- B. applicable to practice.
- C. not based on research.
- D. scientifically supported.

Example 9: Sub-skill 27

Which of the following is the best title for this passage?

- A. The US foreign policy
- B. Lessons from the war in Iraq
- C. The sleepy superpower awakes
- D. The US empire and the new challenges