International Journal of English Language Education
ISSN 2325-0887
2013, Vol. 1, No. 3

The Relative Difficulty and Significance of Reading Skills

Gholamreza Hessamy

Assistant professor, Dept. of English Language and Literature

Payame Noor University, Lashkarak Road, Tehran, Iran

E-mail: yhessamy@yahoo.com

Somayeh Sadeghi (Corresponding author)

Payame Noor University, Lashkarak Road, Tehran, Iran

E-mail: sadeghi.somayeh@yahoo.com

Received: February 4, 2013 Accepted: July 16, 2013 Published: July 17, 2013

doi:10.5296/ijele.v1i3.4017 URL: http://dx.doi.org/10.5296/ijele.v1i3.4017

Abstract

The importance of reading in English as a foreign language has resulted in numerous studies on the construct of reading and the nature of its underlying skills. This study investigated the possibility of any hierarchical relationship among a set of hypothesized reading comprehension (RC) subskills measured by IELTS in terms of two important criteria of relative difficulty and significance. Furthermore, it examined the relationship between the level of difficulty and significance of the subskills. To this end, five reading subskills that are most frequently referred to in the literature were selected and presented to six expert judges in order to assign them to IELTS reading items. The results showed considerable agreement among judges on matching test items with the subskills. Then, the selected items were administered to 180 Iranian English majors. Analysis of data, using repeated measure ANOVA, showed significant differences among mean scores on four, out of five, reading subskills in terms of difficulty; therefore, a hierarchy of difficulty was identified. In addition, using structural equation modeling, contribution of each of the five subskills to the construct of reading ability turned out to be significant indicating a hierarchy of relative contribution to overall RC ability among the specified subskills. However, Spearman rank order correlation showed no significant relationship between difficulty level and relative contribution to the totality of RC. The findings have implications for test construction as well as teaching and material development.

Keywords: reading subskills, significance of reading subskills, difficulty of reading subskills, IELTS reading test

1. Introduction

The amount of research conducted on the nature of reading is enormous. Among different views on the nature of reading, one way to analyze the complex process of reading is dividing reading comprehension into a set of subskills. Some studies that investigated the issue identified a single factor (e.g. Rost, 1993; Spearritt, 1972; Drahozal & Hanna, 1979; Carver, 1992; Lunzer et al, 1979; Rosenshine, 1980). Based on such studies, advocates of the unitary view argue that reading is an indivisible, holistic skill that can be simply called reading ability. There have also been qualitative studies supporting the unitary view including Alderson and Lukmani (1989) and Alderson (1990) which reported difficulty in getting agreement among judges about the subskills being tested by particular items. However, the validity of Alderson's qualitative study was questioned by Matthews (1990), Weir et al (1990), and Weir and Porter (1994).

On the other hand, other studies concluded that reading can be divided into a set of distinct subskills. Alderson (2000) cites quantitative studies supporting multi-divisible view such as Carroll (1993), Drum et al (1981), Pollitt et al (1985), and Davey (1988). Qualitative studies against a unitary view of reading include Anderson et al (1991), Bachman et al (1988), Brutten et al (1991), Lumley (1993), Weakly (1993) and Weir et al (1990), as cited by Weir and Porter (1994). In support of multi-divisible view of reading, Lumley (1993) argues that it may not be possible to say that reading subskills exist objectively as separate, concrete elements. However, the fact that the literature is full of reference to them and that they are often considered in syllabus design and test construction suggests that they are inherently appealing as a working construct to many teachers and testers. Besides, Hughes (2003) holds that the fact that not all of the subskills involved in reading have been confirmed by research is not a justified reason to exclude them from our test specifications. He holds that if we want to test overall reading ability, a sample of the subskills involved in reading in relation to our purpose should be included in the test.

2. Reading Sub-skills in Standard Proficiency Tests

The concept of reading subskills is widely employed as a basis for constructing standard language proficiency tests. The influence of subskills is evident in TOEFL, IELTS, FCE, and TEEP. Matthews (1990) believes that existing taxonomies are mainly derived from the Munby's and are frequently used in standard tests of reading (e.g. IELTS).

Likewise, Lumley (1993) refers to Carroll (1980) who identified language skills as part of test content specification and holds that Munby's taxonomy of language skills has been one of the most productive sources of information for the purpose of test construction. Carroll (1980) considers eleven subskills of the Munby's list as appropriate for testing.

Weir (1988) examining the development of TEEP tried to include discrete assessment of the reading subskills in addition to integrated performance tasks, as noted by Lumley (1993). He offers a list of reading comprehension subskills relevant to EAP contexts.

Alderson (1990a) also argues that TEEP and ELTS are clearly based on a taxonomic view of language skills. He claims that both tests can be related to explicit statements about the

subskill measured by each test item. He considers Weir (1983) as the source of these statements for TEEP and Criper & Davis (1988) for ELTS.

Reading components were also referred to in the FCE Handbook (1997) when it describes a reading test as expecting candidates to be able "…to show understanding of gist, detail and text structure and to deduce meaning" (FCE Handbook, 1997:7), or when it specifies each of the four parts included in the test as measuring a particular subskill of reading: the first part measures main idea, the second part is concerned with both detailed and global understanding, the third part involves inferring meaning and lexical reference, and the last part measures the ability of locating information in different sections of text.

3. Hierarchies of Reading Sub-skills

Studies that investigated reading subskills examined such issues as how many distinguishable skills are involved in reading; whether judges can assign particular subskills to particular test items; whether some subskills presuppose others; which subskills are more demanding; and what possible relationships exist among them. Various lists and taxonomies of subskills were developed as a result of these studies, including Gray (1960), Barrett (1968), Davis (1968), Lunzer et al (1979), and Munby (1978). Attempts have been made in a number of these studies to arrange subskills as hierarchies. Urquhart and Weir (1998) believe that Lunzer et al (1979) taxonomy is organized as a hierarchy, with the 'lowest level' skill at the top. Munby's taxonomy, they assert, was not intended to be hierarchically arranged, yet they refer to Mead's (1982) review of the work in which he "argued that it should have been, on the grounds that some skills seem to presuppose the learning of other skills" (p. 91). Matthews (1990) also believes that Munby's list represents a hierarchy, arguing that the list is so arranged that the smaller linguistic units precede the larger ones and, therefore, contains strings of implicationally related categories. Weir (1983) considers his list of skills an 'ordered list' and "claims that the enabling skills tested by TEEP are ordered such that lower order skills are required before higher order skills can be deployed" (Alderson, 1990:429).

However, Alderson and Lukmani (1989) and Alderson (1990) examining the idea of a hierarchy of reading sub-skills according to level of cognitive ability found little agreement among judges on the cognitive level of most items. Weir et al (1990) criticizing the insufficient definitions of the terms 'high' and 'low' order questions in Alderson's (1990a) study claim that it is not surprising that the teachers participating in the study changed their minds over level of the questions.

The level of operation is not the only criteria for a hierarchy. Difficulty has been also used as a criterion for ranking subskills into a hierarchy. Urquhart and Weir (1998) refer to Lunzer et al examining the pragmatic validity of their hierarchy of skills and argue that they failed to find evidence for this but found some evidence of a hierarchy of difficulty.

Alderson and Lukmani (1989) carried out an item analysis of performance of students on the items which judges agreed were testing particular subskills at lower or higher levels of difficulty. The results demonstrated little relationship between item difficulty and item level. Also, Alderson (1990) by examining student performance on TEEP and ELTS items found

little relationship between level of test items –'higher' or 'lower order'– and their actual difficulty for test takers. Weir et al (1990) criticize Alderson's view of the implicational relationship between 'high' and 'low' order questions and the assumption he made about the relative difficulty of high and low order questions – that items assumed to measure higher-order skills must be more difficult than those assumed to measure lower-order ones – and argue that there is no reason to expect that the cognitive level of skills tested by two items determine the relative difficulty of those items.

Studies which identified hierarchies of difficulty of subskills include Clymer (1968), as noted by Hudson (2007), who believes the categories in his taxonomy are arranged from easy to difficult in the following order: literal comprehension, reorganization, inferential comprehension, evaluation, and appreciation. Mecartty (1998) also identified three subskills of locating details, simple inferential skills, and complex inferential skills and examined the possibility of identifying a hierarchy for them. The results indicated a hierarchy according to which the third subskill was more difficult than the first and the second ones. Likewise, Lumley (1993) arrived at a very considerable agreement among judges on the difficulty level of the subskills in his study as well as a high correlation between statistical evidence on the difficulty of the items and the judges' consensus on their difficulty level.

Although reading subskills have been investigated in many studies, it is still far from clear how they are related to one another. This study examined the relationships that may exist among reading comprehension sub-skills in terms of difficulty and significance. More specifically, the following research questions were investigated:

1. Is it possible to identify a hierarchy of difficulty among reading subskills?

2. Is it possible to identify a hierarchy of relative contribution to the overall reading ability?

3. Is there any significant relationship between difficulty level of reading subskills and their relative importance?

4. Method

4.1 Participants

Two groups of subjects took part in this study: expert judges and university students.

The first one was a group of six English teachers who were asked to make judgments about the subskills tested by each item in an IELTS academic reading test. Four of them had MA degrees and two had BA degrees in TEFL. The expert judges had enough experience in teaching IELTS preparation courses.

The second group consisted of 180 junior and senior English majors from whom the data for this study was obtained. They were students of TEFL, English Literature, and English translation. Due to their educational background, they were considered as a homogeneous group, especially with regard to their level of proficiency. All the participants had the same native language background, i.e. Persian and were between 21 and 34 years old.

4.2 The choice of Subskills

First, lists of skills suggested by Davis (1968), Barrett (1968), and Lunzer et al (1979) were examined. The given descriptions either seemed too general and undefined or appeared not to be so relevant to the purpose of this study which focused mainly on subskills measured by the IELTS reading test. In addition, as pointed out by Lumley (1993), in most cases, these studies were based on the development of reading in children who acquire English as their first language. As it is not yet clear whether adults' second/foreign language learning involves the same processes as children acquiring or learning language skills in their mother tongue, these taxonomies were considered to be inappropriate for this study. Likewise, the review of Munby (1978) revealed that not only there are some overlaps between the skills, but also they are poorly distinguished. As a result, it was decided to develop a new list including those subskills that are most frequently mentioned in the literature. Thus, the following five subskills which were believed to be highly distinguished and explicitly defined were selected:

1- Identifying main idea

2- Understanding specific information

3- Making inferences

4- Identifying writer's views/claims

5- Extracting information from a prose text to put into diagrammatic representation

It should be noted that subskills 1, 2, 4 and 5 are often mentioned in IELTS test specifications as well.

4.3 Materials

Two instruments were used in this study. First, a checklist was used to determine the subskills that are tested by each item in an IELTS academic reading test. A list of five reading subskills was given to the expert judges who were asked to read through the accompanying test paper and select, from the list, what the main, highest level subskill required to answer each item, i.e. the subskill without which it would not be possible to answer the item, is.

The second instrument was an IELTS academic reading test used to elicit information on the relative significance and difficulty level of the selected subskills. The original test consisting of 40 items and three reading passages would take 60 minutes. After the experts having judged the subskill tested by each item, 21 items based on two passages were selected as the final test to be answered in 33 minutes. The items selected were those on which a high level of agreement was achieved with regard to measuring a particular subskill. The internal consistency of the 21 items selected as the final test calculated through Cronbach's Alpha was .75.

5. Results

High levels of agreement were achieved in matching IELTS reading items to the five specified subskills.

Table 1. Ascribing Items to Subskills by Judges

Item	Subskill	Degree of agreement (percent)
1-5	2	100
6-9	4	100
10-12	1,2,3,4	< 50
13-18	1	83
19-24	4	100
25-27	2,3	66
28-32	4	100
33,34,36,37	2	83
35	2,3	66
38-40	5	100

As it can be seen in Table 1. almost all experts agreed on which subskill was being tested by 23 items, i.e. more than half of the items. The experts showed a good level of agreement in recognizing the nature of the ability that was intended to be tapped by each item.

With regard to the first question of the study, in order to suggest a hierarchy of difficulty among the subskills, a repeated measures ANOVA was run to compare the mean scores on each of the five subskills.

Table 2. Repeated Measures ANOVA for Reading Comprehension Subskills

Effect		Value	F	df	Error df	Sig.	Partial Eta Squared
Tests	Pillai's Trace	.588	62.876[a]	4	176	.000	.588
	Wilks' Lambda	.412	62.876[a]	4	176	.000	.588
	Hotelling's Trace	1.429	62.876[a]	4	176	.000	.588
	Roy's Largest Root	1.429	62.876[a]	4	176	.000	.588

As Table 2. shows, the F-observed value (62.87) is greater than the critical value of F at 4 and 176 degrees of freedom. Therefore, it can be concluded that there are significant differences

among the mean scores on different reading subskills; hence, it can be claimed that a hierarchy of difficulty among the subskills exists.

Table 3. Descriptive Statistics for Reading Comprehension Sub-skills

	Mean	Std. Deviation	N
Identifying writer's views/claims	3.39	1.085	180
Understanding specific information	2.99	1.262	180
Identifying main idea	2.72	1.442	180
Making inferences	2.25	1.657	180
Extracting information from a text to put into diagrammatic representation	1.67	1.615	180

According to Table 3., "Identifying the writer's views/claims" is the easiest reading subskill with a mean of 3.39, while "Extracting information from a text to put into diagrammatic representation" with a mean of 1.67 is the most difficult one. Table 3 shows the relative difficulty of the subskills based on the mean scores.

Table 4. Post-Hoc Comparison of Reading Comprehension Sub-Skills

(I) Tests	(J) Tests	Mean Difference (I-J)	Std. Error	Sig.	95% confidence interval for difference	
					Lower Bound	Upper Bound
Identifying Writer's Views/ claims	Understanding Specific Information	.394*	.101	.001	.107	.682
	Identifying Main Idea	.667*	.104	.000	.370	.963

	Making Inferences	1.139*	.124	.000	.787	1.491
	Extracting Information	1.722*	.110	.000	1.409	2.035
Understanding Specific Information	Identifying Main Idea	.272	.118	.221	-.063	.607
	Making Inferences	.744*	.139	.000	.348	1.140
	Extracting Information	1.328*	.128	.000	.963	1.693
Identifying Main Idea	Making Inferences	.472*	.132	.004	.097	.848
	Extracting Information	1.056*	.121	.000	.710	1.401
Making Inferences	Extracting Information	.583*	.129	.000	.217	.950
*. The mean difference is significant at the .05 level.						

As shown in Table 4., post-hoc comparison tests were employed to compare the means in pairs in order to identify the order of difficulty of the reading subskills. Based on the results displayed in Table 4, it can be concluded that:

1. “Identifying the writer’s views / claims” is the easiest subskill with a mean score of 3.39. This figure shows a difference that is statistically significant when compared with the other four subskills. The asterisks placed on the mean scores indicate significant differences between the two means compared.

2. There is not any significant difference between the mean scores of subskills 2 and 3. That is to say, the difference between “Understanding specific information” (2.99) and “Identifying main idea” (2.72) is not significant. However, “Understanding specific information” with a mean score of 2.99 is significantly different from the other two subskills.

3. “Identifying main idea” with a mean score of 2.72 is significantly different from “Making inferences” and “Extracting information from a text to put into diagrammatic representation”.

4. The fourth sub-skill, “Making inferences” with a mean score of 2.25 has a statistically significant difference from the mean of “Extracting information from a text to put into diagrammatic representation”.

Based on the results of the post-hoc comparisons applied in this study, the following order of subskills from the easiest to the most difficult will appear:

1. Identifying the writer’s views/claims

2. Understanding specific information and Identifying main ideas

3. Making inferences

4. Extracting information

To answer the second research question concerning the relationship among reading comprehension subskills in terms of significance, Structural Equation Modeling (SEM) was used. The five subskills examined in this study were believed to tap the underlying construct of L2 reading comprehension ability.

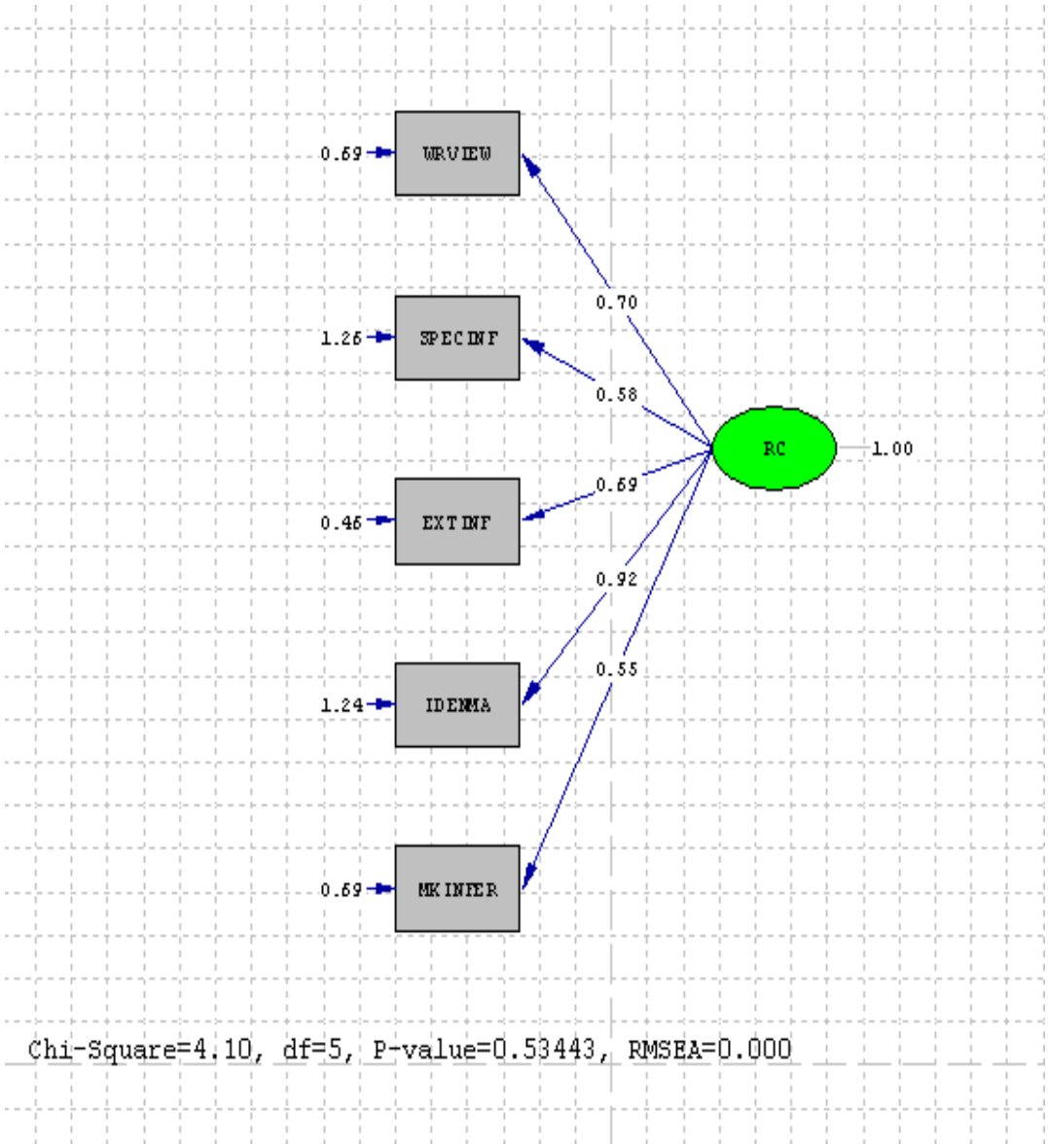

Figure 1. The structural model of L2 reading subskills

Based on the APA considerations, all researches that include structural models should provide the following information together with the model (Schumacker and Lomax, 2004)

- The Covariance matrix of the data: Table 4

- The GFI (goodness of fit index): .99
- The AGFI (Adjusted goodness of fit index): .97
- NFI (Normed fit index): .98

Note: GFI, AGFIU and NFI should be equal to or higher than .90 to accept the model as a fit one.

The produced structural model (Figure 1.) shows the five subskills and their contributions to the hypothetical latent variable of L2 reading. The model enjoys a good fit. The chi-square value of 4.10 has a P-value of .53 (a good model is one with a P-value of higher than .05). On the other hand, the Root Mean Square of Error Approximation of zero (0.000) is another proof of the model fit. This index should be lower than .05 for a good model.

The coefficients connecting the oval shape – which represents the latent variable of L2 reading comprehension – to the five subskills can be viewed as regression weights. Based on these structural weights the following order of importance can be suggested:

Table 5. Covariance matrix of RC subskills

	WRVIEW	SPECINF	EXTINF	IDEMA	MKINFER
WRVIEW	1.18				
SPECINF	.47	1.59			
EXTINF	.48	.37	.94		
IDEMA	.65	.59	.61	2.08	
MKINFER	.35	.25	.42	.51	.99

1. Identifying main idea (IDEMA) (with the greatest contribution to the totality of RC)
2. Identifying the writer's views/claims (WRVIEW)
3. Extracting information from a text to put into diagrammatic representation (EXTINF)
4. Understanding specific information (SPCINF)
5. Making inferences (MKINFER)

The third research question investigated the relationship between the relative difficulty of the subskills and their relative importance.

Table 6. Order of Difficulty and Importance

	Difficulty	Importance
WRVIEW	4	2
SPECINF	3	4
EXTINF	1	3
IDEMA	3	1
MKINFER	2	5

Table 6. presents the relative difficulty and significance of the five subskills of this study (rank 1 represents the most significant or the most difficult subskill).

Table 7. Spearman Rank-Order Correlation

		Importance
Difficulty	Correlation Coefficient	-.462
	Sig. (2-tailed)	.434
	N	5

Spearman rank-order correlation was run to probe the relationship between the order of difficulty and importance of the RC subskills. The Spearman Rho turned out to be -.46 (P = .434 > .05) as shown in Table 7. Since the Rho value is lower than the critical value of 1, it can be concluded that there is no significant relationship between difficulty level of reading subskills and their importance.

6. Discussion

This study showed that experts are able to agree on the subskills involved in answering particular test items. This is in line with the findings of Weir et al (1990), Anderson et al. (1991), Brutten et al. (1991), Lumley (1993), Weakly (1993), and Bachman et al. (1988) who achieved a sufficient degree of agreement on what are being measured by individual test items, but contrary to the findings of Alderson and Lukmani (1989) and Alderson (1990a) who concluded that judges are not able to agree on assigning particular subskills to individual test items.

Attempts were made in this study to specify the subskills which were explicitly definable and caused little ambiguity. This was perhaps one of the reasons for the judges' considerable agreement on the task of matching subskills with test items. Another reason could be the judges' experience in IELTS preparation courses which led to their familiarity with the skills tested in IELTS and caused them little difficulty in identifying the nature of the items. In the meantime, it should be noted that the descriptions of some of the subskills in the study were

somehow related to the tasks involved in the test, which is not surprising, of course. Naturally, this could have also contributed to the judges' substantial agreement on assigning particular test items to the reading subskills.

On the other hand, the findings on the relationships between different components of reading in terms of difficulty were like Lumley (1993) who fit reading subskills "into broad bands of increasing difficulty" (p.48). In addition, the results arranging reading skill components into a hierarchy based on their relative difficulty support the conclusion of Clymer (1968) with the categories in his taxonomy being ordered from easy to difficult according to their demand, as well as Mecartty's (1998) findings indicating that "complex inferential skills" was more difficult than two other subskills being investigated. It should be noted that the subskills identified in Clymer (1968) and Mecartty (1998) were different from those examined in this study. However, all these findings contradict those of Alderson and Lukmani (1989) and Alderson (1990a) who were not able to find a hierarchy of difficulty of subskills agreed upon as higher and lower order.

Meanwhile, the findings of this study have to some extent answered the question posed by Weir and porter (1994) and Urquhart and Weir (1998) concerning the relative importance of reading subskills. The latter argued that further research could contribute to the rejection of a fully unitary view of reading by establishing "the extent to which each component has a meaningful effect on the measurement of reading comprehension" (p. 14). The subskills investigated in this study proved to have a significant contribution to the measurement of L2 reading. Each component explained 55 to 92 percent of the overall variance in the test. The findings of this study lend support to the idea of multi-divisibility of L2 reading ability and possibility of identifying separate reading sub-skills as held by Dublin et al. (1986), William and Moran (1989), Carr and Levy (1990), and Grabe (1991).

Finally, although previous research has rarely addressed the issue of the relationship between the relative difficulty and relative importance of reading subskills, the findings of this study which revealed the absence of any significant relationship can raise new questions regarding the crucial question of how difficulty and importance of subskills may interact. Although it may seem logical to assume that the more important a reading component is the more demanding or difficult it might be, the findings of the present study can cast doubt on this assumption. It seems that further research is absolutely needed as the nature of the relationship between these two factors can have valuable implications for teaching and testing reading.

7. Conclusion

The focus of this study was on the relationships that may exist among L2 reading subskills in terms of difficulty and significance. It was successful in achieving high levels of agreement among judges on determining the skills involved in answering test items. This could be encouraging for a componential view of reading that considers reading as divisible into a set of components measurable by particular test items.

With regard to the first and second research questions, significant differences were found

among reading subskills in terms of difficulty and significance. Based on the findings, it can be concluded that it is possible to suggest empirically-based hierarchies of difficulty and importance among the subskills. This can be taken as evidence that the selected subskills exist separately, and thus, they can be used for writing items. That is, it is possible to develop items specifically measuring these subskills in a reading test. In addition, knowing the difficulty level of different subskills helps test developers to make their test more suitable to test takers at a particular level of difficulty. Furthermore, awareness of the relative importance of reading subskills can clarify which ones have priority to be included in a test so that it would be a more appropriate measure of test takers' reading ability. Test developers can design items measuring the subskills which have a more significant role in demonstrating reading ability. All these insights can enhance the theory-based validity of reading tests.

Moreover, the findings provide some useful feedback to teaching reading. Teachers and course designers, especially those involved in test preparation courses for IELTS and other standard tests of proficiency would be able to devote more time and effort to practicing those subskills which are more difficult and more significant in their contribution to the totality of reading ability and reading test performance.

Of course, further research is needed to study more reading skills in different contexts, using different instruments with perhaps more participants in order to shed further light on the nature of L2 reading ability.

References

Alderson, J.C. (1990). Testing reading comprehension skills (Part One). *Reading in a Foreign Language*, *6*(2), 425-438.

Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.

Alderson, J. C., & Lukmani, Y. (1989). Cognition and reading: cognitive levels as embodied in test questions. *Reading in a Foreign Language*, *5*(2), 253-270.

Anderson, N. J., Bachman, L., Perkins, K., and Cohen, A. (1991). An exploratory study into the construct validity of a reading comprehension test: Triangulation of data sources. *Language Testing*, *8*(1), 41-66.

Bachman, L. F., Vanniarjan, A. K. S., & Lynch, B. (1988). Task and ability analysis as a basis for examining content and construct comparability in two EFL proficiency batteries. *Language Testing*, *5*(2), 128-159.

Barrett, T. C. (1968). What is reading? Some current concepts. In H. M. Robinson (Ed.), *Innovation and change in reading instruction: The sixteenth handbook of the National Society for the Study of Education*. Chicago: The University of Chicago Press.

Brutten, S. R., Perkins, K., and Upshur, J. A. (1991). Measuring growth in ESL reading. Paper presented at the 13th Annual Language Testing Research Colloquium, Princeton, New Jersey.

Carr, T. H., & Levy, B. A. (Eds.). (1990). *Reading and Its Development: Component Skills*

Approaches. San Diego: Academic Press.

Carroll, B. J. (1980). *Testing Communicative Competence; An Interim Study*. Oxford: Pergamon Press.

Carroll, J. B. (1993). *Human Cognitive Abilities*. Cambridge: Cambridge University Press.

Carver, R. P. (1992). What do standardized tests of reading comprehension measure in terms of efficiency, accuracy and rate? *Reading Research Quarterly*, *27*, 347-359.

Clymer, T. (1968). What is 'reading'?: Some current concepts. In H. M. Robinson (Ed.), *Innovation and change in reading instruction* (pp. 7-29). Chicago: The National Society for the Study of Education.

Criper, C., & Davis, A. (Eds.) (1988). ELTS validation project report. *English Language Testing Service Research report* Vol. 1(i). The British Council and University of Cambridge Local Examinations Syndicate.

Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly*, *3*, 499-545.

Davey, B. (1988). Factors affecting the difficulty of reading comprehension items for successful and unsuccessful readers. *Experimental Education*, *56*, 67-76.

de Witt, R. (1997). *How to Prepare for IELTS*. London: The British Council.

Drahozal, E.C. and Hanna, G. S. (1979). Reading comprehension subscores: pretty bottles for ordinary wine. *Journal of Reading*, *21*(5), 416-20.

Drum, P. A., Calfee, R. C., and Cook, L. K. (1981). The effects of variables on performance in reading comprehension tests. *Reading Research Quarterly*, *16*, 486-514.

Dublin, F., Eskey, D. E., & Grabe, W. (Eds.) (1986). *Teaching Second Language Reading for Academic Purposes*. USA: Addison-Wesley Publishing Company.

Grabe, W. (1991). Current developments in second language reading. *TESOL Quarterly*, *25*(3), 375-396.

Gray, W. S. (1960). The major aspects of reading. In H. Robinson (Ed.), *Sequential developments of reading abilities, no 90* (pp. 8-24). Chicago: Chicago University Press.

Hudson, T. (2007). *Teaching Second Language Reading*. London: Oxford University Press.

Hughes, A. (2003). *Testing for Language Teachers* (2nd ed.). Cambridge: Cambridge University Press.

Lumley, T. (1993). The notion of sub-skills in reading comprehension tests: an EAP example. *Language Testing*, *10*(3), 211-234.

Lunzer, E., Waite, M., and Dolan, T. (1979). Comprehension and comprehension tests. In E. Lunzer, & K. Gardner, editors, *The effective use of reading*. London: Heinemann Educational Books.

Matthews, M. (1990). Skill taxonomies and problems for the testing of reading. *Reading in a Foreign Language*, *7*(1), 511-517.

Mead, R. (1982). Review of Munby, J. 'Communicative syllabus design'. *Applied Linguistics*, *3*(1), 70-77.

Mecartty, F. H. (1998). The effects of proficiency level and passage content on reading skills assessment. *Foreign Language Annals*, *31*(4), 517-534.

Munby, J. (1978). *Communicative Syllabus Design*. Cambridge: Cambridge University Press.

Pollitt, A., Hutchinson, C., Entwistle, N., & DeLuca, C. (1985). *What Makes Exam Questions Difficult? An Analysis of 'O' Grade Questions and Answers*. Edinburgh: Scottish Academic Press.

Rosenshine, B. V. (1980). Skills hierarchies in reading comprehension. In Spiro et al, editors, *Theoretical issues on language testing* (pp. 535-554). Hillsdale, NJ: Erlbaum.

Rost, D. (1993). Assessing the different components of reading comprehension: fact or fiction? *Language Testing*, *10*(1), 79-92.

Spearritt, D. (1972). Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Reading Research Quarterly*, *8,* 92-111.

Urquhart, A. H., & Weir, C. J. (1998). *Reading in a Second Language*. London: Longman.

Weakly, S. (1993). Procedures in the content validation of an EAP proficiency test of reading comprehension. *Unpublished MA thesis*. University of Reading.

Weir, C. J. (1983). Identifying the language needs of overseas students in tertiary education in United Kingdom. *Unpublished PhD dissertation*. University of London.

Weir, C. (1988). The specification, realization and validation of an English language proficiency test. In A. Hughes (Ed.), *Testing English for university study, ELT documents*, *127,* 45-110.

Weir, C. J., Hughes, A., & Porter, D. (1990). Reading skills: Hierarchies, implicational relationships and identifiability. *Reading in a Foreign Language*, *7*(1), 505-510.

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*, *10*(2), 1-19.

Williams, E., & Moran, C. (1989). Reading in a foreign language at intermediate and advanced levels with particular reference to English. *Language Teaching*, *22*(4), 217- 228.

Copyright Disclaimer