

## **Factors Related to the Difficulty of College Scholastic Ability Test English Items**

**Kyung-Ae Jin** (Korea Institute of Curriculum & Evaluation)\*

**Chung Park** (Korea Institute of Curriculum & Evaluation)

**Jin, Kyung-Ae; & Park, Chung. (2005). Factors related to the difficulty of college scholastic ability test English items. *English Teaching*, 60(3), 17-25.**

This study has attempted to explore the factors related with the difficulty of the College Scholastic Ability Test (CSAT) English items. The test used for the present study was 2005 English section of CSAT, which was composed of 50 items. The subjects were 570,431 students who took 2005 CSAT English test. 16 independent variables were correlated with the proportion of the students' correct responses (dependent variable) in 50 items. The findings support the previous study with regard to the nature of correlation between the item or text variables and proportion of the correct answers in the 2003 CSAT English test (Kyung-Ae Jin & Chung Park, 2004). Also, it was found that the 16 variables suggested as the factors related with the item difficulty in this study explain approximately 52% ( $R^2= 0.519$ ) of the variance of the proportion of the students' correct responses in the 2005 CSAT English test. This indicates that the variables identified in the present study are rather explanatory in terms of predicting CSAT English item difficulty.

### **I. INTRODUCTION**

Item difficulty in high stakes test such as College Scholastic Ability Test (CSAT) in Korea is one of the critical matters to concern when developing the test. Even though CSAT grade reports provide standard and percentile scores, while not providing raw scores, maintaining appropriate difficulty levels of the test each year is important. It is, however, difficult to predict item difficulty of the test in advance since many factors are related to the difficulty of an item. In order to explore the factors affecting difficulty of CSAT English items, Kyung-Ae Jin and Chung Park (2004) have attempted correlational analysis between the variables that were expected to affect difficulty of an item and the proportion of correct responses. The data used in this previous study were 2003 CSAT English items and the students' response in each item. 2003 CSAT English section has 50 items and out of 50, 17

---

\* 1<sup>st</sup> author: Kyung-Ae Jin, 2<sup>nd</sup> author: Chung Park

items were listening comprehension items. The total number of subjects were 355,146.

The study found that among the 46 independent variables, 14 variables were identified as having meaningful correlation with item difficulty, and the regression analysis with the 14 variables explained 49% of the factors influencing difficulty of an item. The independent variables which were found to be related with the 'higher proportion of correct responses' — easier item — were 'inferring assertion', 'inferring the author's emotional status', 'understand topic', 'understand atmosphere', 'understand assertion', 'number of text sentences', 'main idea first sentence: Main idea information is in first sentence of the text', and 'information in the last sentence'.

On the other hand, the variables which were found to be related with the 'lower proportion of correct items' — more difficult item — were 'inferring cause', 'inferring the parenthesis', 'summarize the paragraph', 'grammar', 'text sentence words: average number of words per text sentence', and 'number of semicolons, colons and dashes used in text' (Kyung-Ae Jin & Chung Park, 2004). This study, however, was done only with the students' responses of 2003 CSAT, there exists a need to retest the independent variables with those of other year's CSAT for generalization of the results.

Upon this need, based on the previous studies with regard to the variables that affect the CSAT English item difficulty (Kyung-Ae Jin & Chung Park, 2004) and other studies (Freedle & Kostin, 1987; Kostin, 2004), the present study attempts to explore the factors related to the difficulty of 2005 CSAT English items.

## II. RESEARCH DESIGN

### 1. Data

The data used in this study was 2005 CSAT English items, which was implemented to the students in November, 2004, and the students' responses in each item. 2005 CSAT English section has 50 items, and out of 50, 17 items were listening comprehension items. The item difficulty index used in the study was 'proportion of correct responses out of the number of the total responses' for each item. The total number of students who took 2005 CSAT English section were 570,431. The reliability index of 2005 CSAT English test was .92 in Cronbach  $\alpha$  and this indicates the test was very reliable.

### 2. CSAT (2005') English Item Analysis

In order to investigate the variables affecting item difficulty of CSAT, English items of 2005 CSAT were analyzed with regard to 'item variables', and 'text variables' and

‘text/item overlap variables’. Item variables were specified based on the test specifications of CSAT. Text variables were analyzed based on the linguistic nature of the text. Text/item overlap variable is the one that was drawn with regard to the interaction between text and the item.

### 1) Item Variables

The item variables were categorized as five domains, which corresponds to the ‘Performance Domain’ specification in CSAT English section, and those are ‘literal understanding’, ‘inference’, ‘main idea’, ‘application’, ‘grammar’ and ‘vocabulary’. ‘Literal understanding’ asks the literal or factual information of the passage. The items in ‘inferences’ domain were supposed to measure the students’ ability to infer the following aspects; ‘inferring assertion of the text’, ‘inferring emotional status’, ‘inferring omitted part’, ‘inferring referent of the text’, and ‘inferring object/incident/relationship’.

The variables under ‘Main Idea’ domain, which asks global understanding of the passage has been categorized into several, but ‘understanding topic’ variable was the only one that was significantly related with the item difficulty in 2003 CSAT. Therefore, in the present study, in the ‘main idea’ domain, ‘understanding topic’ variable’ will be analyzed as an independent variable.

The ‘Application’ domain includes the items designed for indirect assessment of speaking and writing skill. As the national foreign language curriculum includes speaking and writing components, CSAT includes the items measuring speaking and writing ability indirectly. The typical speaking item is the items that ask the most appropriate expression in the context after hearing the listening passage. On the other hand, the writing items include summarizing the

**TABLE 1**  
**Item Variables (2005’ CSAT English)**

Domain	Variables	Number of Items	2005 CSAT Item Number
1. Literal understanding	V1. literal understanding	11	1,4,5,7,9,10,11,12,32,33,34
	V2. inferring assertion	5	6,39,40,45,47
2. Inferences	V3. inferring emotional status	3	2,38,44
	V4. inferring omitted part	4	26,27,28,29
	V5. inferring referent	3	19,46,50
	V6. inferring object/incident/ relationship	4	3,8,13,18
3. Main idea	V7. understanding topic	5	25,30,31,42,43
	V8. summarizing	2	36,37
4. Application	V9. sentence/paragraph order	3	35,41,49
	V10. appropriate expression	4	14,15,16,17
5. Grammar	V11. Grammar	3	20,21,22
6. Vocabulary	V12. Vocabulary	2	23,24,48

passage or finding out the relevant order of sentences considering the coherence. In the 'Application' domain, 'appropriate expression', 'summarizing', 'sentence/paragraph order' variables will be analyzed. The other domains are 'Grammar' and 'Vocabulary'. The 2005 item variables, the number of items which belong to each item variable and the item numbers are presented in the Table 1.

## 2) Text Variables and Text/Item Overlap Variables

The text variables and text/item overlap variables which were assumed to be related with the item difficulty of the 2005 CSAT English test were 't1: number of sentences of the text', 't2: average number of words per text sentence', 't3: number of semicolons, colons and dashes used in the text', 't4: relevant information in the last sentence of text'. These text and text/overlap variables are ones that were found to be related with the item difficulty in the previous study (Kyung-Ae Jin & Chung Park, 2004). The assumption underlying text variables is that as the frequency of 'number of sentences of the text', 'average number of words per text sentence', 'number of semicolons, colons and dashes used in the text', in the text of each item increase, the items will be more difficult. On the other hand, when the 'relevant information in the last sentence of text' — text/item overlap variable — exists, the item becomes easier.

## 3. Data Analysis

In order to figure out the factors related to the item difficulty of CSAT English section, the correlation analysis between the sixteen independent variables — twelve item variables and four text variables and test/item overlap variables — and the proportion of correct answers of 570,431 students was made. In addition, the regression analysis has been implemented to provide the explanatory power of the independent variables. The computer program used for the analysis was SPSS version 11.0 (statistical package for social science).

# III. RESULTS

## 1. The Results of the Correlation Analysis

The independent variables—both item variables and the text and text/item overlap variables and the proportion of the 570,431 students' correct answers on the 50 items were correlated. The results are in Table 3. Among variables, v1 ~ v12 are item variables, t1 ~ t3

are text variables and t4 is text/item overlap variable.

According to Table 3, v11:grammar and v12:vocabulary show significant negative correlation with the proportion of the lower correct answers. In other words, in the 2005 CSAT English test, grammar and vocabulary item are significantly related with the lower proportion of the correct answers. 'v4: inferring omitted part' variable and the text variable 't2: average number of words per text sentence' are also negatively correlated with the lower proportion of the correct answers, therefore, affect the item difficult.

**TABLE 3**  
**The Correlational Coefficients between Independent Variables and**  
**the Item Difficulty (2005' CSAT English)**

Variables	Correlation Coefficient
v1. literal understanding	.158
v2. inferring assertion	-.033
v3. inferring emotional status	.204
v4. inferring omitted part	-.153
v5. inferring referent	.133
v6. inferring object/incident/relationship	.103
v7. understanding topic	.058
v8. summarizing	-.004
v9. sentence/paragraph order	.145
v10. appropriate expression	.087
v11. grammar	-.415**
v12. vocabulary	-.414**
T1. number of sentences of the text	.077
t2. average number of words per text sentence	-.171
t3. number of semicolons, colons and dashes used in the text	.051
T4. relevant information in the last sentence of text'	.275

\*\* : p<.05

On the other hand, the factors that are positively correlated with the proportion of the students' correct answers are, 'v1:literal understanding', 'v3: inferring emotional status' and 't4:relevant information in the last sentence of text'. In other words, item variables such as 'literal understanding', 'inferring emotional status' and the text/item overlap variable, 'relevant information in the last sentence of text' are significantly related with the easiness of the item.

## 2. The Results of the Regression Analysis

The regression analysis has been made and the results are in Table 5 below. According to the regression analysis, v4 (-2.063\*): 'inferring omitted part', v11 (-3.726\*\*): 'grammar', and v12 (-3.543\*\*): 'vocabulary' affect the item difficulty significantly ( $\alpha = 0.05$ ). Also, it

was found that the 16 variables suggested as the factors related with the item difficulty in this study explain approximately 52% ( $R^2 = 0.519$ ) of the variance of the proportion of the students' correct responses in the 2005 CSAT English test. This indicates that the variables identified in the present study have explanatory power in terms of predicting CSAT English item difficulty.

**TABLE 5**  
**The Results of Regression Analysis (2005' CSAT English)**

Model	variable	B	S.E.of B	T	R <sup>2</sup>
	constant	78.720	12.033	6.542	
2005 CSAT English	V2	-9.524	9.004	-.159	-.058
	V3	-.233	12.434	-.003	-.019
	V4	-20.059	9.725	-.303	-2.063**
	V5	6.770	10.089	.090	.671
	V6	-1.397	9.029	-.021	-.155
	V7	-8.622	9.322	-.144	-.925
	V8	-9.639	12.042	-.105	-.800
	V9	7.554	10.382	.100	.728
	V10	-2.943	9.791	-.044	-.301
	V11	-39.908	10.712	-.528	-3.726**
	V12	-36.888	10.413	-.488	-3.543**
	T1	-.965	.636	-.264	-1.518
	T2	-.310	.491	-.107	-.631
	T3	-.517	1.431	-.051	-.361
T4	7.070	8.053	.144	.878	

\*\* :  $p < .05$

#### IV. DISCUSSION

This study has attempted to explore the factors related with the difficulty of the CSAT English items. The test used for the present study was 2005 English section of CSAT, which was composed of 50 items. The subjects were 570,431 students. 16 independent variables, which were assumed to be related to the item difficulty were correlated with the proportion of the students' correct responses (dependent variable) on 50 items.

Among the 16 variables, the variables found to be positively related to (related to the easiness of the item) the item difficulty were, 'v1:literal understanding', 'v3: inferring emotional status' and 't4:relevant information in the last sentence of text'.

On the other hand, v11:grammar and v12:vocabulary show significant negative

correlation with the proportion of the lower correct answers. In other words, in the 2005 CSAT English test, grammar and vocabulary item are significantly related with the lower proportion of the correct answers. 'v4: inferring omitted part' variable and the text variable 't2: average number of words per text sentence' are also negatively correlated with the lower proportion of the correct answers, therefore, cause the item difficult.

These findings support the previous study on the correlation between the item or text variables with proportion of the correct answers in the 2003 CSAT English test. According to Kyung-Ae Jin and Chung Park (2004), "If the relevant information is in the last sentence of the text, this makes the item easier. Also, with regard to the item variables, inferring cause, inferring the phrases omitted, grammar items are found to be difficult. With regard to the text variables, when the length of the sentence is longer — in other words, if average number of words per text sentence is greater — items become more difficult."

According to the regression analysis, v4(-2.063\*): 'inferring omitted part', v11 (-3.726\*\*): 'grammar', and v12(-3.543\*\*): 'vocabulary' affect the item difficulty significantly ( $\alpha = 0.05$ ). Also, it was found that the 16 variables suggested as the factors related with the item difficulty in this study explain approximately 52% ( $R^2 = 0.519$ ) of the variance of the proportion of the students' correct responses in the 2005 CSAT English test. This indicates that the variables identified in the present study are rather explanatory in terms of predicting CSAT English item difficulty. Further and continuous studies on the factors related with English item difficulty are required in order to provide more information to the item developers.

## REFERENCES

- Drum, P. A., Calfee, R. C., & Cook, L. K. (1981). The effects of surface structure variables on performance in reading comprehension tests. *Reading Research Quarterly*, 16, 486-514.
- Elder, C. (1992). How do subject specialists construe second language proficiency? *Melbourne Papers in Language Testing*, 1(1), 16-33.
- Elder, C. (1993). *The Proficiency Test for Language Teachers (Italian)*: Final report on the test development process. Melbourne: Language Testing Research Centre.
- Elder, C. (1994). Performance testing as benchmark for loter teacher education. *Melbourne Papers in Language Testing*, 3(1), 1-25.
- Ellis, R. (1984). *Classroom second language development*. Oxford: Pergamon Press.
- Freedle, R., & Kostin, I. (1991). *The prediction of SAT reading comprehension item difficulty for expository prose passages* (ETS Research report RR-91-29). Princeton, NJ: Educational Testing Service.

- Freedle, R., & Kostin, I. (1992). *The prediction of GRE reading comprehension item difficulty for expository prose passages for each of three item types: Main ideas, inferences, and explicit statements* (ETS Research report RR-91-59). Princeton, NJ: Educational Testing Service.
- Freedle, R., Kostin, I., & Schwartz, L. (1987). *A comparison of strategies used by black and white students in solving SAT verbal analogies using a thinking aloud method and a matched percentage-correct design* (ETS Research Report RR-87-48). Princeton, NJ: Educational Testing Service.
- Freedle, R., Fine, J., & Fellbaum, C. (1981). *Predictors of good and bad essays*. Paper presented at the annual Georgetown University Roundtable on languages and linguistics, Washington, D. C.
- Hare, V., Rabinowitz, M., & Schieble, K. (1989). Text effects on main idea comprehension. *Reading Research Quarterly*, 24, 72-88.
- Jin, Kyung-Ae; & Park, Chung. (2004). The prediction of english item difficulty in college scholastic ability Test. *English Teaching*, 59(1), 267-277.
- Jin, Kyung-Ae; Lee, Yi-Gap; Chang, Kyung-Sook; Lee Hoi-Joo; & Cho, Ji-Min. (2002). *Analysis of 2003 Preliminary College Scholastic Ability Test -English-*. CAT 2002-25: Seoul: Korea Institute of Curriculum & Evaluation.
- Kaufman, D., & Brooks, J. G. (1996). Interdisciplinary collaboration in teacher education: A constructivist approach. *Tesol Quarterly*, 30(2), 231-251.
- Kostin, I. (2004). *Exploring item characteristics that are related to the difficulty of toefl dialogue items* (ETS Research Report RR-04-11). Princeton, NJ: Educational Testing Service.
- Long, M., & Porter, P. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19(2), 207-228.
- McNamara, T. F. (1996). *Measuring second language performance*. London and New York: Addison Wesley Longman.

Applicable levels : Secondary

Key words: College Scholastic Ability Test, Item difficulty

Kyung-Ae Jin  
Division of College Scholastic Ability Test  
Korea Institute of Curriculum & Evaluation  
25-1, Jongro-Gu Samchung-Dong  
Seoul, Korea  
Tel: 02-3704-3653  
Fax: 02-3704-3540  
E-mail: kajin@kice.re.kr

Chung Park  
Division of Educational Evaluation  
Korea Institute of Curriculum & Evaluation  
25-1, Jongro-Gu Samchung-Dong  
Seoul, Korea  
Tel: 02-3704-3586  
Fax: 02-3704-3540  
E-mail: chungpark@kice.re.kr

Received in May, 2005  
Reviewed in June, 2005  
Revised version received in August, 2005