한국수학교육학회지 시리즈 A 〈수학교육〉
1993. 6. 제 32권 제 3호 244-255

Journal of the Korea Society of Mathematical Education
Series A : June 1993, Vol. 32. No. 3, 244-255

한미 대학 입학 시험(수학)의 비교 연구:
7차 실험 평가와 S.A.T.를 중심으로

Korea-U.S. Cross-National Comparison Study on Mathematics College Entrance Exams : the 7-th Pilot Test and the S.A.T.*

Kwon, Oh Nam**(Indiana Univ.)

INTRODUCTION

This study is motivated by the changes in the Scholastic Aptitude Test and reform in the *College Entrance Exam in Korea*. The redesigned S.A.T. in the United States will be administered in Spring 1994 and the new version of the *College Entrance Exam in Korea* will be started in August 1993. The three main changes in the S.A.T.-Mathematics in the United States are (1) permission to use calculators, (2) inclusion of open-ended questions, and (3) content revisions consistent with the National Council of Teachers of Mathematics *Curriculum and Evaluation Standards* (Braswell, 1992). On the other hand, the two main changes in the *College Entrance Exam of Korea*, in general, are (1) change from estimating on numerous subjects to four core subjects (Korean, mathematics, physical science and social science, and English) with emphasis on interconnectedness of subjects, and (2) the exam can be taken twice a year instead of only once (Report on pilot test of *College Entrance Exam*, 1992). The changes in the Korean mathematics subtest are (1) greater emphasis on problem solving which requires higher-order thinking, and (2) more-realistic problems in which mathematics is applied to other subject areas. Up to this point, 7 pilot tests have been completed. The validity of the S.A.T. as a measure of college aptitude has been discussed by several writers (Boldt, 1986; Cowen, 1991; Wilmouth, 1991). However the validity of the new *Korean College Entrance Exam* has not been discussed. Since the goals of both tests are similar, the students scores on the S.A.T. and 7th pilot test of the *Korean Exam* will be correlated to examine the validity of the *College Entrance Exam*. [1]The following questions were investigated:

1. Are there statistically significant sex-related differences in the means of the pilot test-M and the S.A.T.-M?

---

* The research reported in this paper is part of the author's master thesis completed at Indiana University under the direction of Dr. Peter Kloosterman. The author would like to thank Drs Peter Kloosterman, Frank K. Lester Jr., and Diana Lambdin Kroll for their valuable comments and suggestions.

** Department of Mathematics and Department of Curriculum and Instruction, Indiana University, Bloomington, Indiana 47405, U. S. A.

2. Are there statistically significant differences between groups of students with different intended majors in the pilot test-M and the S.A.T.-M?

3. Do the S.A.T.-M and the 7-th pilot test-M measure the same capability?

SAMPLE

The subjects used for this investigation were 283 senior high school students from three schools in Kwangju, Korea. Two schools are girls' high schools and the other one is a boy's high school. Three classes are composed of students whose intended majors in college are physical science, engineering, and medicine. The other three classes are composed of students whose intended majors in college are literature, social studies, language, and law. The schools each served a student population with a wide range of socioeconomic backgrounds. Kwangju, located in the southwestern part of Korea, is the fifth largest city with population more than one million. Senior high school students were selected for the study because they are the first generation affected by the new Korean College Entrance Exam and have taken the 7-th pilot test.

PROCEDURE AND DATA COLLECTION

The actual 1991 S.A.T.-M (Form Code 1Q from 5 SATs, The College Board,1992) was given to the subjects on March 16, 1993 (male), April 1, 1993 (female), and April 13, 1993 (female). The collected data were high school rank, comprehensive achievement tests scores, mathematics scores, the 7-th pilot test-Mathematics subtest scores, the S.A.T.-Mathematics. The comprehensive achievement tests consist of tests for 16 subjects. Mathematics is a subtest of comprehensive achievement tests. The 7-th pilot test was taken by subjects on Nov. 10, 1992. As Table 1, the 7-th pilot test is composed of Korean language, Mathematical Principles and Scientific Inquiry I-Mathematics, Mathematical principles and Scientific Inquiry II-Science, and Foreign Language (English).

Table 1. Pilot Test for Korean College Entrance Exam

Domain	Number of points	Number of items	Time(min.)	Format
Korean	60	60	90	Written/oral
Math.	40	20	60	Written
Science	60	60	90	Written
English	40	50	70	Written/oral
Total	200	190	310	

DATA ANALYSIS

Mean Scores of the S.A.T.-M and the 7-th pilot test-M

This section presents the results of the analyses of the 7-the pilot test-M and the

S.A.T.-M. Data from the top 30% students were selected for analyses because these students are more likely to be accepted by colleges. In Table 3, the term major 1 is used to indicate that the students' intended major in college are related to social studies, language, law, music, and fine arts(i.e., non-science major). Major 2 represents the students' intended majors in college are physical science, mathematics, medicine, and engineering (i.e., science major). Major 1 students are taking mathematics I and major 2 students are taking mathematics II which includes the materials in mathematics I. The means and standard deviations of the 7-th pilot test-M are given in Table 2-3. The only significance result is F-ratio of 4.95 compared by major.

Table 2. Mean(Mode) Scores of the 7-th Pilot Test-M Compared by Gender

	Male n=98		Female n=185		F (1,83) (1,281)
	Mean	S.D.	Mean	S.D.	
Top30%	14.62(18.0)	5.35	13.68(14.0)	4.59	.72
Total	11.63(8.00)	4.94	11.44(8.00)	4.52	. 11

Table 3. Mean(Mode) Scores of the 7-th Pilot Test-M Compared by Major

	Major1 n=140		Major2 n=143		F (1,83) (1,281)
	Mean	S.D.	Mean	S.D.	
Top30%	13.10(8.00)	4.79	14.89(14.0)	4.81	.29
Total	10.89(8.00)	4.37	12.11(12.0)	4.88	4.95*

*p<.01

Tables 4-5 report the results of ANOVA with the means and standard deviations of the S.A.T.-M. This indicates that there were significant differences between majors and top 30% groups in each major group.

Table 4. Mean(Mode) Scores of the S.A.T.-M Compared by Gender

	Male n=91		Female n=184		F (1,83) (1,274)
	Mean	S.D.	Mean	S.D.	
Top30%	657.93(670)	68.89	637.86(650)	73.40	1.49
Total	579.23(700)	120.79	581.69(610)	95.26	.34

Table 5. Mean(Mode) Scores of the S.A.T.-M Compared by Major

	Major1 n=136		Major2 n=139		F (1,83) (1,274)
	Mean	S.D.	Mean	S.D.	
Top30%	615.29(610)	73.30	673.49(650)	58.68	16.40*
Total	545.30(610)	103.98	615.68(670)	92.23	36.60**

*p<.005
**p<.0001

Since there are essentially four different groups, which are major1 & male, major2 & male, major 1 & female, and major 2 & female, the followings will show their differences. Table 6 shows the means and standard deviations of these four groups.

Table 6. Means for the 7-th Pilot Test-M Compared by Groups

	Male				Female			
	n=47		n=51		n=93		n=92	
	Major1		Major2		Major1		Major2	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Top30%	15.29	5.74	14.00	5.07	12.00	3.89	15.36	4.68
Total	10.98	5.14	12.24	4.73	10.84	3.95	12.04	4.99

Table 7 reports the results of the analysis of variance between the groups.

Table 7. Anova of Four Groups' the 7-th Pilot-Test Score(Top 30%)

Source	df	S.S.	M.S.	F-ratio	Prob
Between Groups	3	186.71	62.24	2.816	.0442
Within Groups	81	1789.29	22.09		

The F-ratio of 2.816 was significant. This indicates that top 30% students of major 2 & female group are significantly different at the $p<.05$ level than other groups. Table 8 reports the means and standard deviations for the S.A.T.-M of four different groups.

Table 8. Means for the S.A.T.-M Compared by Groups

	Male				Female			
	n=42		n=48		n=93		n=91	
	Major1		Major2		Major1		Major2	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Top30%	640.00	83.02	674.67	49.69	602.86	66.04	672.86	63.82
Total	549.07	129.89	606.25	106.22	543.55	90.30	620.66	84.13

Table 9 AND 10 report the results of the analysis of variance between four groups.

Table 9. Anova of Groups' the S.A.T-M Score

Source	df	S.S.	M.S.	F-ratio	Prob
Between Groups	3	348013.28	116004.43	11.972	.0000
Within Groups	271	2625977.26	9689.95		

Table 10. Anova of the Top 30% Groups' the S.A.T.-M Score

Source	df	S.S.	M.S.	F-ratio	Prob
Between Groups	3	85001.46	28333.82	6.522	.0005
Within Groups	81	251916.19	4344.64		

This indicates that major 2 groups are significantly different from major 1 groups among overall and the top 30% students.

The mean scores of the S.A.T.-M between two countries are compared in Table 11. Females' mean scores are lower than males' mean scores in the U.S.A. But this data shows that there is no significant gender difference in Korea. The difference between the two countries is above 100 points. Hollway(1988) compared data from Japan and the United States. Some of the major findings from this report indicate that effort is believed to be primary importance in determining achievement in Japan, but ability is seen as the primary factor in the United States.

Table 11. Mean Scores of the S.A.T.-M Compared by Nationality

	U.S.A. n=1,032,685			KOREA n=275	
	Mean	S.D.		Mean	S.D.
Male	497	127	Male	579	121
Female	453	115	Female	582	95

Mathematics as intended major	605		Major2	616	92
Total	474	123	Total	581	104

Correlations between comprehensive achievement test scores, mathematics scores, the 7-th pilot scores, and the S.A.T.-M scores

This section of the chapter presents the correlations between the following variables: comprehensive achievement scores, mathematics scores, the 7-the pilot-M scores, and the S.A.T.-M scores. Table 12 reports the correlations compared by gender and major.

Table 12. Correlations between Comprehensive Achieve Tests, Mathematics, the 7-th Pilot Test, and the S.A.T. Scores Compared by Gender and Major

<Major1>

	Male 43≤n≤47			Female n=93			Male + Female 136≤n≤140		
	MATH	PILOT	SAT	MATH	PILOT	SAT	MATH	PILOT	SAT
COM	.90**	.69**	.72**	.69**	.19	.52**	.80**	.43**	.60**
MATH		.74**	.73**		.26*	.53**		.45**	.60**
PILOT			.66**			.18			.41**

<Major2>

	Male 48≤n≤51			Female 91≤n≤92			Male + Female 139≤n≤143		
COM	.83**	.50**	.58**	.67**	.35**	.62**	.76**	.38**	.59**
MATH		.57**	.70**		.60**	.52**		.56**	.59**
PILOT			.50**			.34**			.34**

<Major1 + Major2>

	Male 91≤n≤98			Female 184≤n≤185		
COM	.85**	.60**[1]	.64**	.67**	.28**[1]	.54**
MATH		.63**[2]	.65**[3]		.44**[2]	.44**[3]
PILOT			.60**[4]			.29**[4]

[1,4] significant differences ($p < .01$) between the correlations

[2,3] significant differences ($p < .05$) between the correlations

* significant at the .05 probability level

** significant at the .01 probability level

Sex-related differences in correlations were computed using Fisher's r to Z-transformation (Hays,1973). As can be seen in Table 12, there were four statistically significant sex-related differences in the correlations. At p=.01 level, the correlations between COM & PILOT, and PILOT & SAT were significantly stronger for males than for females. At p=.05 level, the correlations between MATH & PILOT, and MATH & SAT were significantly stronger for males than for females. There were no statistically major-related significant differences between the correlations. The COM & PILOT, and PILOT & SAT were the only non-significant correlations for the major 1 female subjects.

Table 13. Correlations between Comprehensive Achieve Tests, Mathematics, the 7-th Pilot Test, and the S.A.T. Scores for the Top 30% Students

<Major1>

	Male n=14				Female n=28				Male + Female n=42		
	MATH	PILOT	SAT		MATH	PILOT	SAT		MATH	PILOT	SAT
COM	.88**	.76**	.69**	COM	.43**	.17	.07	COM	.62**	.53**	.40**
MATH		.81**	.67**	MATH		.20	.54*	MATH		.47**	.59**
PILOT			.66**	PILOT			.04	PILOT			.38**

<Major2>

	Male n=15				Female n=27				Male + Female n=42		
COM	.82**	.71**	.73**	COM	.41**	-.08	.55*	COM	.54**	.28	.58**
MATH		.63*	.61*	MATH		.39*	.49*	MATH		.49	.51**
PILOT			.57*	PILOT			.23	PILOT			.32*

<Major1 + Major2>

	Male n=29				Female n=55		
COM	.85**	.74**	.66**	COM	.44**	.11	.32*
MATH		.74**	.61**	MATH		.35*	.54**
PILOT			.56**	PILOT			.29*

* significant at the .05 probability level

** significant at the .01 probability level

DISCUSSION

While there was evidence of sex-related differences in the S.A.T.-M in the U.S.A. (see also Table 10), there were no significant sex-related differences in the S.A.T.-M and the 7-th pilot test-M. Many writers indicated that performance on mathematics (also higher-order thinking tasks) is affected by student's interest, beliefs, confidence, perseverance, and ability to monitor their own progress (Kloosterman, 1988 & 1991; Lester & Kroll, 1990). In all large-scale assessment, two of the most powerful predictors of student achievement have been increased time on mathematics and the taking of advanced coursework. The NAEP documents the substantial growth in achievement that is associated with advanced course taking. Similarly, Myers and Milne (1983, 1988) reported that course taking was the single most powerful predictor of mathematics achievement. This is in agreement with the finding of the study which shows that the major 2 subjects were significantly stronger than the major 1 subjects in the mean scores for the 7-th pilot test-M and the S.A.T.-M (Table 3 & Table 5). Yet, there were somewhat contradictory findings with respect to gender difference between the correlations. The correlations between measured variables for the female subjects were significantly lower than those for the male subjects.

CONCLUSION

1. There existed no significant differences in the S.A.T.-M and the 7-th pilot test-M between the female and male students. Sex-differences in mathematics ability has been discussed by many writers (Benbow, 1988; Silver et al., 1988). Hanna and Kündiger (1986) indicated that there were sex-differences for grade 12 students from the 15 countries which participated the Second International Mathematics Study (SIMS). However, the current study demonstrated that the female subjects were not significantly lower than males in the 7-the pilot test-M and S.A.T.-M scores. This finding does not agree with the above results. Yet, analyses by Ethington (1990) of data for eight countries which participated in the SIMS suggest that gender differences were small, not uniformly in favor of males, and inconsistent across countries.

2. Achievement on mathematics is affected by the student's attitude towards mathematics, and course-taking. There were strong significant intended major-related differences in the S.A.T.-M and the 7-th pilot test-M. This result is more likely attributed to the type of students who will select their major as science related or non-science related. The major 2 subjects
are more likely to have positive attitudes towards mathematics than the major 1 subjects since they chose science related careers. Furthermore, the major 2 subjects are taking more advanced course (mathematics-II) than the major 1 subjects. This result agrees with the results of the 1986 NAEP (Dossey et al., 1988) and the 1978 NAEP achievement data (Welch et al., 1982).

3. The content validity of two tests is different. The correlations between two tests were lower than the correlations between several other variables. In addition, the comprehensive achievement tests were highly correlated with the S.A.T.-M. This indicated that the

characteristics of the S.A.T.-M might be toward aptitude rather than achievement. The new S.A.T. starting March 1994 will be of the same difficulty as the current ones (The College Board, 1992). It is speculated that the new S.A.T. and the new *Korean College Entrance Exam* have different content validity in mathematics.

4. The validity of the pilot test-M is questioned for the major 1 female subjects. The correlation between COM and PILOT for this group(r=.1877) was not statistically significant (r=.1877). For the top 30% students of the same group, the comprehensive achievement test score was not significantly correlated with the pilot test-M (r=.1094). Therefore, the fairness of the pilot test-M is questioned for the major 1 female subjects. In general, the correlations between measured variables for the females were lower than those of the males. The possible explanation for this phenomenon is related to the findings of Hacket and Betz (1989). They reported that both mathematics performance and self-efficacy were positively and significantly correlated with attitudes towards mathematics and masculine sex-role orientation. They also concluded that differences in females' and males' mathematics self-efficacy expectations were correlated with performance differences between the two groups.

FUTURE DIRECTIONS

This research study compared the new Korean College Entrance Exam and the S.A.T. As a result of conducting this experiment, the investigator recommends the consideration of the following items for future research based on the results and conclusions of the study.

1. A longitudinal study of the new Korean College Entrance Exam should be conducted to determine the validity and appropriateness of this test in making selection for admission to college.

2. Comparison studies needs for the new Korean College Entrance Exam and the new S.A.T. starting March 1994.

3. An investigation into the factors that explain why the major 1 female subjects had equivalent mean scores on the 7th pilot test-M and the S.A.T.-M and lower correlations with other measured variables than the major 1 male subjects should be warranted.

REFERENCES

Benbow, C. P. (1988), Sex Differences in Mathematical Reasoning Ability in Intellectually Talented Preadolescents: Their Nature, Effects, and Possible Causes, *Behavioral and Brain Sciences*, Vol. 11, pp. 169-183.

Boldt, Robert F. (1986), *Generalization of the S.A.T. Validity across Colleges*, College Entrance Examination Board.

Braswell, James S. (1992), Changes in the S.A.T. in 1994, *The Mathematic Teacher*, Vol. 85, No. 1, pp. 16-21.

Cowen, Sheila and Fion, Sandra J. (1991), *Appropriateness of the S.A.T. in selecting students for Admission to California State University*, Hayward, Paper presented at the Annual meeting of the California Educational Research Association.

Dossey, J. A., Mullis, I. V. S., Lindquist, M. M., and Chamber, D. L. (1988), *The Mathematics Report Card: Are We Measuring Up?* (Trends and Achievement Based on the 1986 National Assessment, Report No. 17-M-01), Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service.

Educational Testing Service(1993), Sample of the new S. A. T.

Ethington, C. A. (1990), Gender Differences in Mathematics: An International Perspective, *Journal for Research in Mathematics Education*, Vol. 15, pp. 361-377.

Hacket, G., & Betz, N. E. (1989), An exploration of the mathematics self-efficacy/mathematics performance correspondence, *Journal for Research in Mathematics Education*, 20, 261-273.

Hanna, G. and Kündiger, E. (1986, April), *Differences in Mathematical Achievement Levels and in Attitudes for Girls and Boys in Twenty Countries*, Paper Delivered at the Annual Meeting of the American Educational Research Association, San Francisco.

Hays, William L. (1973), *Statistics for the Social Sciences*, Holt Rinehart Winston.

Holloway, S. C. (1988), Concepts of ability and Effort in Japan and the United States, Review of Educational Research, Vol. 58, pp. 327-345.

Kloosterman, Peter (1988), Self-Confidence and Motivation in Mathematics, *Journal of Educational Psychology*, Vol. 80, pp. 345-351.

Kloosterman, Peter (1991), Beliefs and Achievement in Seventh-Grade Mathematics, (Focus on

Learning Problems in Mathematics), Vol. 13, No. 3, pp. 3-15.

Lester, Frank K., Jr., and Kroll, Diana Lambdin (1990), Assessing Student Growth in Mathematical Problem Solving, *Assessing Higher Order Thinking in Mathematics*, Washington, DC: American Association for the Advancement of Science.

Myer, D. E. and Milne, A. M. (1983), *Mathematics Achievement of Language-Minority Students*, Paper Prepared for Part C Research Agenda: Math and Language Minority Student Project, Washington, DC: Decision Resources.

Myer, D. E. and Milne, A. M. (1988), Effects of Home Language and Primary Language on Mathematics Achievement, In R. R. Cocking and J. P. Mestre (Eds.), *Linguistic and Cultural Influences on Learning Mathematics*, pp. 259-293.

Report on pilot test of College Entrance Exam (1992), National Educational Assessment Institute (in Korean).

Silver, E. A., Linquist, M. M., Carpenter, T.P., Brown, C. A., Kouba, V. L., & Swafford, J. O. (1988), The Fourth NAEP Mathematics Assessment: Performance trends for instructional indicators, *The Mathematics Teachers, 81*, 720-727.

Welch, W. W., Anderson, R. E., & Harris, L. J.(1982), The effects of schooling on mathematics achievement, *American Educational Research Journal, 19*, 145-152.

Wilmouth, David (1991), *Should the S.A.T. Be a Factor in College Admissions?*, ED345592.

5 SATs, The College Board, 1992

한미 대학 입학 시험(수학)의 비교 연구: 7차 실험 평가와 S.A.T.를 중심으로

본 연구를 수행하게된 동기는 1994년부터 미국에서 S.A.T.를 개정하고, 한국에서는 대학 수학 능력 시험 제도라는 새로운 제도가 도입되는 것에 있다. 대학 학업 적성 평가 제도로서 미국의 S.A.T. 제도에 대한 유효성이 많은 학자들에 의해서 연구되고 있다. 대학 수학 능력 시험과 S.A.T.는 각각 한국과 미국의 대학에서 학업 적성을 측정한다는 면에서 그 목적이 같다. 한국의 대학 수학 능력 시험의 유효성을 연구하기에는 아직 실시되지 않았으므로 너무 이르다고 본다. 대학 수학 능력 시험 제도 확립이 실험 평가에 근거하기 때문에 대학 수학 능력 시험 실험 평가와 S.A.T.를 비교 연구하는 것은 의미가 있다고 본다. 따라서 본 논문에서는 한국의 대학 수학 능력 시험 실험 평가(수리)와 미국의 S.A.T.(수학)와의 상관 관계를 연구한다.

본 연구의 조사 대상으로 선발된 집단으로서 광주시의 3개교 6학급의 고등학교 3학년 283명이 참가하였다.

본 논문에서 다음과 같은 문제가 연구되었다.

1. 7차 실험 평가(수리)와 S.A.T.(수학)의 평균 점수에 대한 남녀 차이의 통계학적 유의성 (statistical significance).
2. 7차 실험 평가(수리)와 S.A.T.(수학)의 평균 점수에 대한 자연계 인문계 차이의 통계학적 유의성 (statistical significance).
3. 한국의 대학 수학 능력 시험 실험 평가(수리)와 미국의 S.A.T.(수학)의 상관 관계.