DOCUMENT RESUME

ED 39? 835 TM 024 604

AUTHOR	Wang, Xiaoping; Maxey, James
TITLE	Gender Differences in the ACT Mathematics Tests: A Cross Cultural Comparison.
PUB DATE	19 Apr 95
NOTE	30p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Francisco, CA, April 19-21, 1995).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Ability; College Entrance Examinations; Comparative Analysis; *Cross Cultural Studies; Foreign Countries; Grade 12; *High School Graduates; High Schools; *Mathematics Tests; Scores; *Sex Differences; *Statistical Distributions; Test Results
IDENTIFIERS	*ACT Assessment; *China

ABSTRACT

This study focused on gender differences in scores on the Americar College Testing Program (ACT) assessment in mathematics for a large sample of U.S. high school graduates (over 1,700,000) and a sample of 321 Chinese twelfth graders. A smaller gender difference was found for the Chinese sample than for the U.S. sample. However, a similar pattern of difference was found for both cultures. Gender differences existed not only in score means but also in score distributions. Gender differences increasingly favored males as subject ability level increased, and the magnitude of gender differences was varied across tasks for both cultures. The homogeneous culture for Chinese males and females might be one cause of the smaller gender difference in the Chinese sample. (Contains 5 tables, 9 figures, and 6 references.) (Author/SLD)

Gender Differences in the ACT Mathematics Tests:

A Cross Cultural Comparison

Xiaoping Wang
Iowa Department of Education

James Maxey
The American College Testing Program

Paper Presented at the 1995 Annual Meeting of NCME
April 19, San Francisco, CA.

Gender Differences in the ACT Mathematics Tests: A Cross Cultural Comparison

Xiaoping Wang
Department of Education
State of Iowa

James Maxey
The American College Testing Program

Abstract

This study focuses on gender differences in the ACT Mathematics scores for a large sample of U.S. high school graduates and a sample of Chinese twelfth graders. A smaller gender difference is found for the Chinese sample than for the U.S. sample. However, a similar pattern of differences is found for both cultures. Gender differences not only exist in score means but also in score distributions. Gender differences increasingly favor males as subjects ability level increases and the magnitude of gender differences is varied across tasks for both cultures. The homogenous culture for Chinese males and females might be one cause of the smaller gender differences in the Chinese sample.

Geometry/Trigonometry (GT). The time limit was 28 minutes for the Chinese sample whereas it was 60 minutes for American graduates.

Subjects

The ACT-M scores and subtest scores were obtained from the two groups (1992 and 1993) of U.S. high school graduates. The total number students in the 1992 data were 832,217 with 374,384 males and 457,833 females whereas the total number students in the 1993 data were 875,603 with 393,707 males and 481,896 females.

The Chinese sample was three hundred and twenty-one twelfth graders who represented Chinese students at two different educational levels (high ability students versus average ability students). Overall 153 male and 168 female Chinese students were involved in this study.

4. Analyses and Results

Both gender differences in score means and distributions are examined in this study. First, the frequencies of the ACT-M scores and three subtest scores are displayed in the Figures 1-8 and Table 1 for the 1992 and 1993 U.S. data. All eight figures show that females have higher frequencies than males in the low score range whereas males have higher frequencies than females in the high score range for both 1992 and 1993 data.

The ACT-M score distributions of Chinese male and female students are displayed in Figure 9. One of the differences from U.S. data is that Chinese males have higher frequencies than females at the two ends (high score and low score) of distribution whereas Chinese females have higher frequencies than males in the middle of the score range.

The Variance Ratio (VR = Variance of males / Variance of females) indicates gender differences in variability between males and females. Some researchers (e.g., Feingold,

1. Introduction

Prior studies indicate that some gender differences do exist in areas such as verbal ability (favoring females), spatial ability (favoring males), and mathematical ability (favoring males). The magnitude of the gender differences in each area has been found to be different across studies. Some researchers (e.g., Linn, 1991) recently reported a decline in the magnitude of gender differences in mathematics. It is important to know what causes the changes in the magnitude of gender differences? Is it because of the real reduction of the gap between males and females or because of the revisions of the test domain, difficulty, formats, and other characteristics? To better understand the real causes of gender differences, two questions need to be answered; first, what is the relationship between task characteristics and gender differences, and second, what is the relationship between persons' ability and gender differences?

2. Purposes

This study will focus on gender differences in the ACT Mathematics scores for a large sample of U.S. students and a sample of students from the People's Republic of China. It attempts to investigate the existence and magnitude of gender differences in ACT Mathematics for U.S. students and Chinese students and to compare the results of two cultures; to explore the relationships between gender differences and subject ability and/or test characteristics; and thereby to understand better the causes of gender differences.

3. Data

Instruments

ACT Assessment is a national standardized test for high school graduates. The ACT Mathematics test (ACT-M) contains three subtests: Pre-Algebra/Elementary Algebra (EA), Intermediate Algebra/Coordinate Geometry (AG), and Plane

1992) reported that males were consistently more variable than females in quantitative reasoning, spatial visualization, spelling, and general knowledge.

In this study, a variance ratio is obtained for each group on the ACT-M and subtest scores (Table 2). Overall, significant variance ratios of the ACT-M scores are found for 1992 (VR = 1.23) and 1993 (VR = 1.9) U.S. graduates. There is a significant variance ratio (VR = 1.36, $p < .01$) for the Chinese sample (Table 3) as well. Similar but slightly smaller variance ratios are also found in the three subtests for both cultures.

Second, effect sizes (ES) are calculated on ACT-M for the three groups (Table 2 & Table 3). Here, effect size, (mean for males - mean for females) / pooled standard deviation, indicates gender difference in mean within each group. A positive ES indicates that males earned higher scores than females, whereas a negative ES means females earned higher scores than males.

In this study, effect size is .26 for 1992 U.S. graduates on ACT-M, .25 for 1993 U.S. graduates, and -.09 for Chinese sample. Overall, the results show that U.S. male students earned higher average scores than U.S. female students in both 1992 and 1993 data, whereas Chinese female students earned slightly higher scores than Chinese male students on the ACT-M test.

Third, since U.S. results show a shift point (ACT-M score of 21) on the male and female ACT-M distribution lines for 1992 and 1993 data, it may be necessary to see the score tendencies below and above the shift point. We divided each year's data into two subgroups with a cut score of 21. A bigger effect size is found for higher score group (ES = .33) than for lower score group (ES = .09) for 1992 data. A similar result was found for 1993 data as well. The effect size for the higher score group is .25 and for the lower score group is .08 (Table 4).

The Chinese results (Table 3) show that for ACT-M, effect size for high ability (ES = .22) students favors males more than for the average ability group (ES = -.26).

In summary, the results of the current study support Feingold's work. Gender differences not only exist in score means but also in score distributions. Among high school students who at least took the ACT Assessment, males are more variable than females in both U.S. samples and Chinese sample. Another cross-culture similarity is that the effect sizes change to favor males as ability level increases.

Given the existence of the interaction between gender and ability level on the ACT-M test, and the existence of gender difference in variability in ACT-M and the subtests for both cultures, it is important to consider group characteristics when reporting gender differences. For example, Benbow conducted several studies on gender differences in mathematically gifted children. Data from one of her studies (Benbow, 1992) demonstrated the interaction between gender and ability level. Her sample of 2,118 students represented the top one percent of students in mathematical ability. On the basis of SAT scores in the 8th grade, each student was placed in one of three groups. Only two of these groups were targeted for study. The high SAT-M group included students in the top quarter of SAT-M scores, whereas the low SAT-M group consisted of students in the bottom quarter of SAT-M scores. Effect sizes were calculated to test the gender and ability interaction (Table 5). Even though all the students were in the top 1% mathematical ability, the effect sizes differed markedly for the top quarter and bottom quarter groups. Effect sizes also increased from 8th grade to end of 12th grade for both the top quarter and bottom quarter groups, and gender differences were .4 SD larger in the top quarter group than in the bottom quarter group.

The results of gender and ability level interaction studies show that researchers need to be cautious in generalizing the results of gender differences from one type of group to other types of groups. If gender differences are modified by ability levels, both primary studies and meta-analysis studies need to control for ability level in future research.

Finally, the interaction of gender and task characteristics is another interest in this study. The effect sizes of three subtests are obtained for three groups in two cultures.

The same pattern of results is found for both cultures. The largest effect sizes are found for the GT subtest (ES = .29 for both 1992 and 1993 U.S. samples, and ES = -.01 for the Chinese sample), the smaller effect sizes are found for the AG subtest (ES = .17 for 1992 U.S. sample, ES = .14 for 1993 U.S. sample, and ES = -.12 for the Chinese sample) and the EA subtest (effect size equals .19 for 1992 U.S. sample, .22 for 1993 U.S. sample, and -.13 for the Chinese sample). If we agree with that the Plane Geometry/Trigonometry (GT) is the most difficult subtest in the ACT-M test, then the findings of this study support Cleary's (1991) report. Cleary said that the female disadvantage increased with age, task complexity, and quantitativeness increase. The findings also support Lohman's (1988) comments. Lohman said, "females excel in retaining and processing information coded in a string format (that is, a format that preserves order), whereas males excel in retaining and processing information coded as images (that is, a format that preserves configual information)." The GT subtest contains items which require more spatial images.

These findings suggest researchers should be cautious about generalizing gender differences from one type of task to other types of tasks. If gender differences are varied across tasks within a domain, both primary and secondary studies need to examine subdomains as well as overall performance in domain.

However, the Chinese data show smaller gender differences than the U.S. data do. Chinese culture tends to provide a more similar environment (e.g., same courses taken and probability of being employed) for males and females as compared to U.S. culture where students have the opportunity to develop more unique ability profiles. The homogenous culture for Chinese males and females might be one cause of the smaller gender differences in the Chinese sample.

REFERENCES

Benbow, C. P. (1992). Academic achievement in mathematics and science of students between ages 13 and 23: Are there differences among students in the top one percent of mathematical ability? Journal of Educational Psychology, 84, 51-61.

Cleary, T. A. (1991). Gender differences in aptitude and achievement test scores. Paper presented at the 1991 ETS Invitational Conference: Sex Equity in Educational Opportunity, Achievement, and Testing, Educational Testing Service, Princeton, New Jersey.

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. Review of Educational Research, 62, 61-84.

Linn, M. C. (1991). Gender differences in educational achievement. Paper presented at the 1991 ETS Invitational Conference: Sex Equity in Educational Opportunity, Achievement, and Testing, Educational Testing Service, Princeton, New Jersey.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), Advances in the Psychology of Human Intelligence (Vol. 4, pp. 181-248). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wang, X. (1993). Gender Differences in Spatial and Verbal Abilities and Mathematics Performance for Chinese Students: A Cross Culture Comparison. Paper presented at the 1993 Annual Meeting of NCME, Atlanta, Georgia.

Table 1. ACTM Score Distributions for the 1992 & 1993 US High School Graduates

	Ideal	Total		Male		Female	
Scores	%	N	%	N	%	N	%
			1992				
31-36	2.4	21676	2.6	14693	3.9	6983	1.5
26-30	13.6	99341	11.9	54.97	14.5	45244	9.9
20-25	34.0	278049	33.4	129412	34.6	148637	32.5
16-19	34.0	288450	34.7	120370	32.2	168080	36.7
14-15	13.6	106491	12.8	41828	11.2	64663	14.1
1-13	2.4	38210	4.6	13984	3.7	24226	5.3
Total	100.0	832217	100.0	374384	100.1	457833	100.0
			1993				
31-36	2.4	24151	2.8	16240	4.1	7911	1.6
26-30	13.6	107493	12.3	58398	14.8	49095	10.2
20-25	34.0	287288	32.8	132374	33.6	154914	32.2
16-19	34.0	309338	35.3	129874	33.0	179464	37.2
14-15	13.6	109264	12.5	42848	10.9	66416	13.8
1-13	2.4	38069	4.4	13973	3.6	24096	5.0
Total	100.0	875603	100.1	393707	100.0	481896	100.0

Table 2. Gender Differences in ACTM and the Subtests for the 1992 & 1993 US Data

	ACTM			ACT-EA			ACT-AG			ACT-GT		
	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female
						1992						
M	20.0	20.7	19.5	10.4	10.7	10.1	10.0	10.3	9.8	10.2	10.6	9.8
SD	4.7	5.0	4.5	3.2	3.3	3.1	2.9	3.0	2.8	2.8	2.9	2.7
ES	.26			.19			.17			.29		
VR	1.23			1.13			1.15			1.15		
						1993						
M	20.1	21.2	19.6	10.4	10.8	10.1	10.1	10.3	9.9	10.2	10.7	9.9
SD	4.8	6.2	4.5	3.2	3.3	3.12	2.9	3.0	2.8	2.8	2.9	2.7
ES	.25			.22			.14			.29		
VR	1.9			1.13			1.15			1.15		

Notes: For the 1992 data, the total number is 832,217 with 374,384 males and 457,833 females.
For the 1993 data, the total number is 875,603 with 393,707 males and 481,896 females.
ES (Effect Size) = (Mean of Males' - Mean of Females') / Pooled Standard Deviation
VR (Ratio of Variances) = Male Variance / Female Variance

Table 3. Gender Differences of ACT Mathematics (Groups by Gender) for the Chinese Sample

Tests	Group	Mean	SD	Cases	Sex	Mean	SD	Cases	ES	VR
ACTEA	High	15.07	1.46	182	M	15.22	1.48	81	+.18	1.36
					F	14.95	1.44	101		
	Ave.	12.53	1.74	137	M	12.27	1.74	71	-.32	
					F	12.82	1.71	66		
ACTAG	High	15.18	1.87	182	M	15.31	1.83	81	+.13	1.22
					F	15.07	1.90	101		
	Ave.	11.77	1.58	137	M	11.56	1.63	71	-.28	
					F	12.00	1.50	66		
ACTGT	High	14.79	1.84	182	M	15.00	1.87	81	+.21	1.24
					F	14.61	1.81	101		
	Ave.	11.75	1.40	137	M	11.73	1.52	71	-.03	
					F	11.77	1.27	66		
ACTM	High	29.07	2.83	182	M	29.41	2.90	81	+.22	1.36
					F	28.79	2.75	101		
	Ave.	23.37	2.50	137	M	23.06	2.59	71	-.26	
					F	23.70	2.37	66		

Table 4. Gender Differences of the Two Groups (High Score vs. Low Score) For the US sample

Groups	ACTM => 21		ACTM < 21	
Sex	Male	Female	Male	Female
		1992		
Mean	25.34	24.28	16.84	16.65
SD	3.47	3.01	2.16	2.20
N	175582	171779	198802	286054
%	46.90	37.52	53.10	63.48
ES	0.33		0.09	
VR	1.33		0.96	
		1993		
Mean	25.19	24.40	16.86	16.69
SD	3.44	3.00	2.13	2.18
N	184366	182141	209341	299756
%	46.83	37.80	53.17	62.20
ES	0.25		0.08	
VR	1.31		0.96	

Notes: ES (Effect Size) = (Mean of Males' - Mean of Females') / Pooled Standard Deviation
VR (Ratio of Variances) = Male Variance / Female Variance

Table 5. SAT-M Scores by Gender for the High School Students Who are in the Top 1% in Mathematical Ability

Sex	Males			Females			
Sample	Mean	SD	N	Mean	SD	N	ES
			Top Quarter				
8th Grade	640	48	367	622	31	100	.40
End High School	747	42	362	714	49	96	.76
			Bottom Quarter				
8th Grade	462	22	248	462	24	282	.00
End High School	631	71	216	613	63	250	.27

Figure 1. 1992 ACTM Score Distributions: Males' vs. Females'

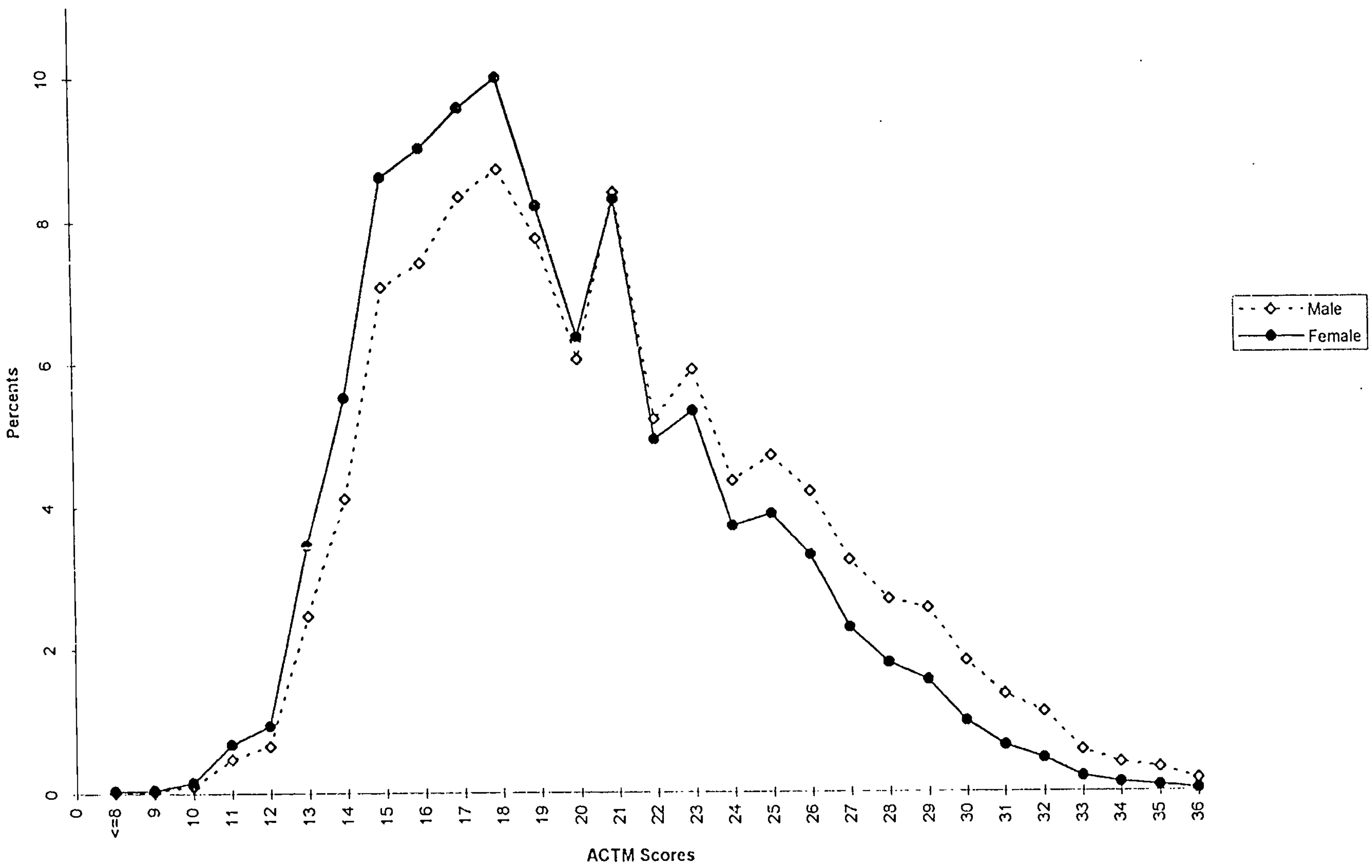
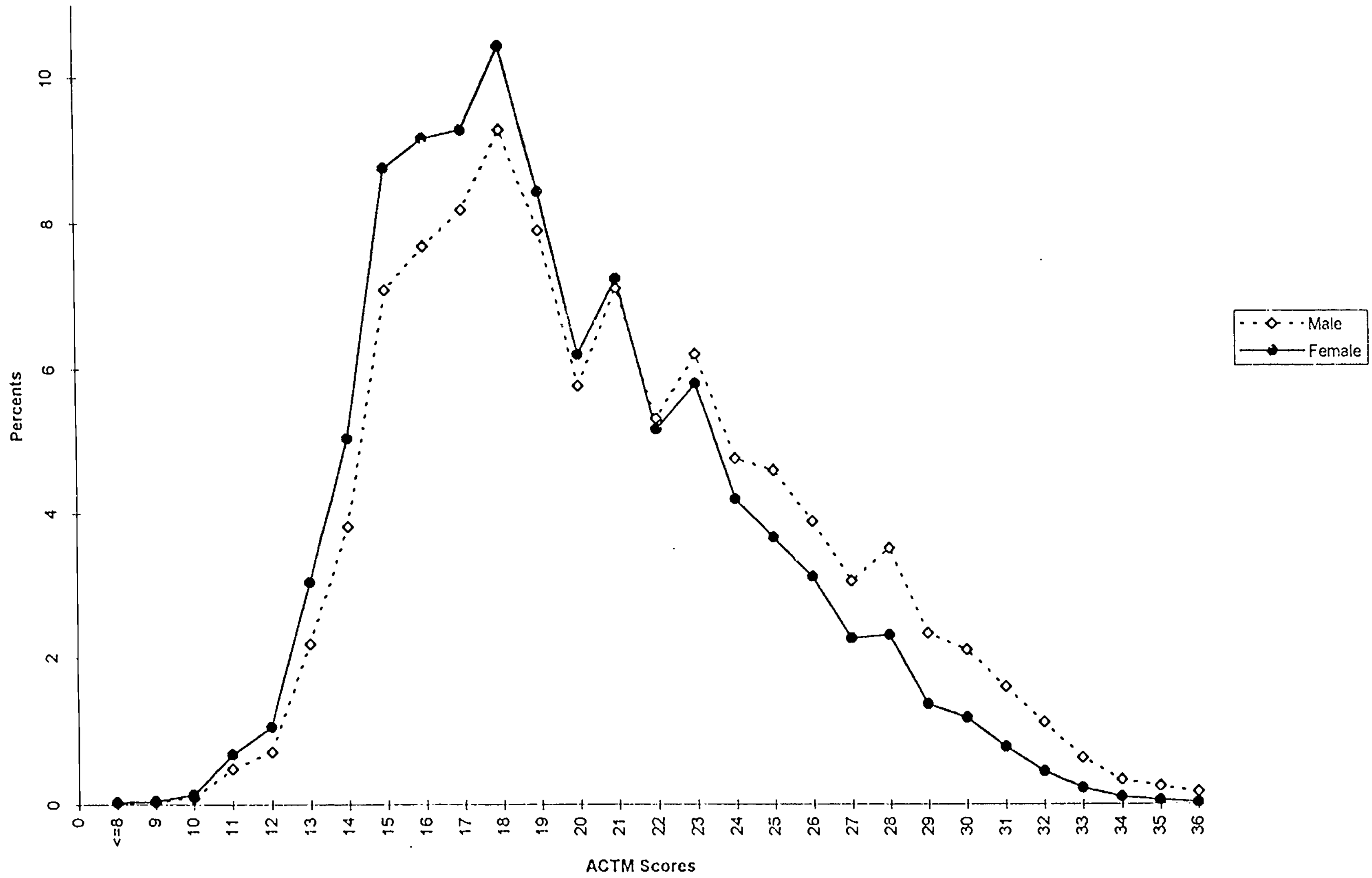


Figure 2. 1993 ACTM Score Distributions: Males' vs. Females'



BEST COPY AVAILABLE

Figure 3. 1992 ACT-EA Score Distributions: Males' vs. Females'

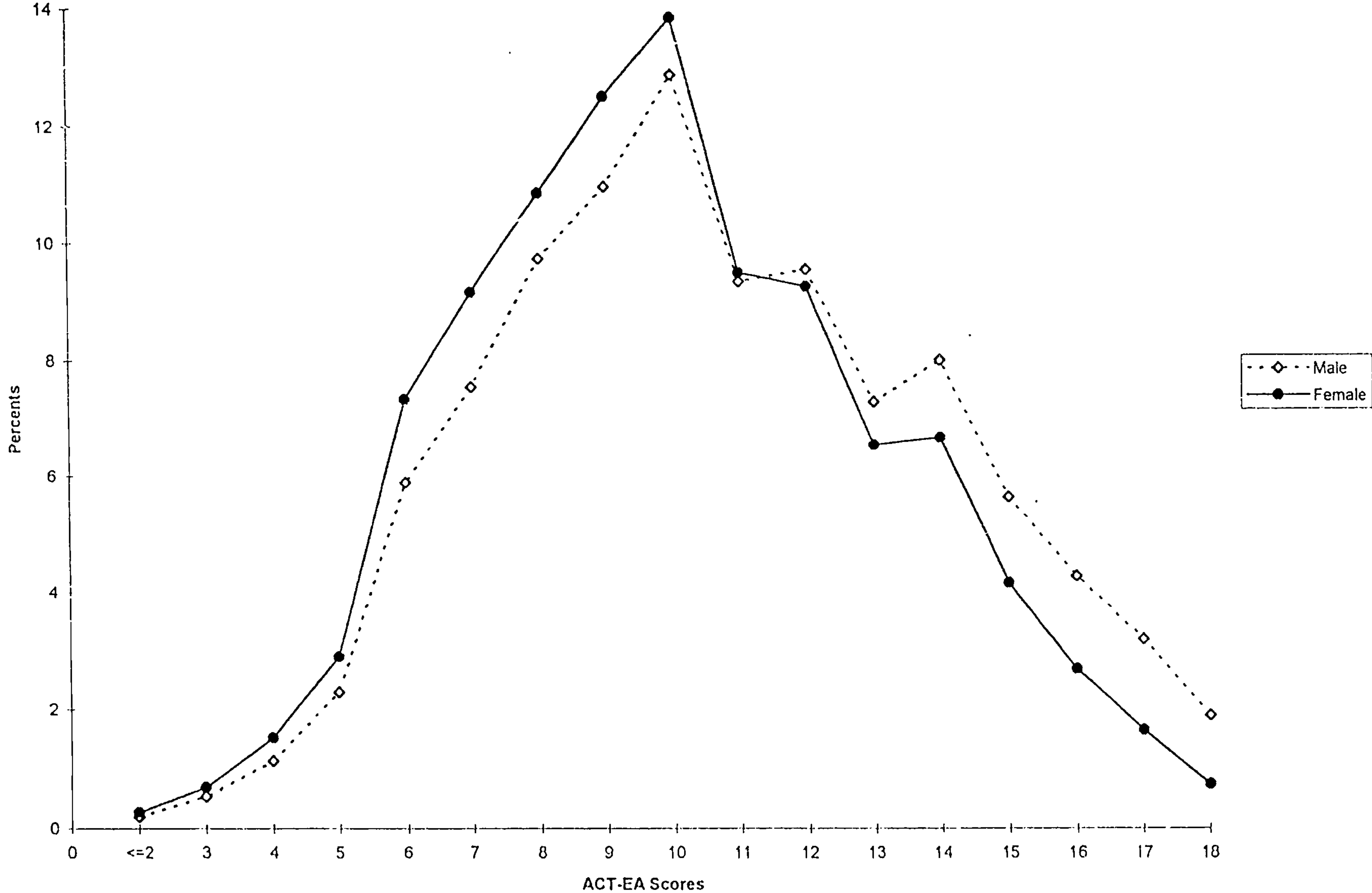


Figure 4. 1993 ACT-EA Score Distributions: Males' vs. Females'

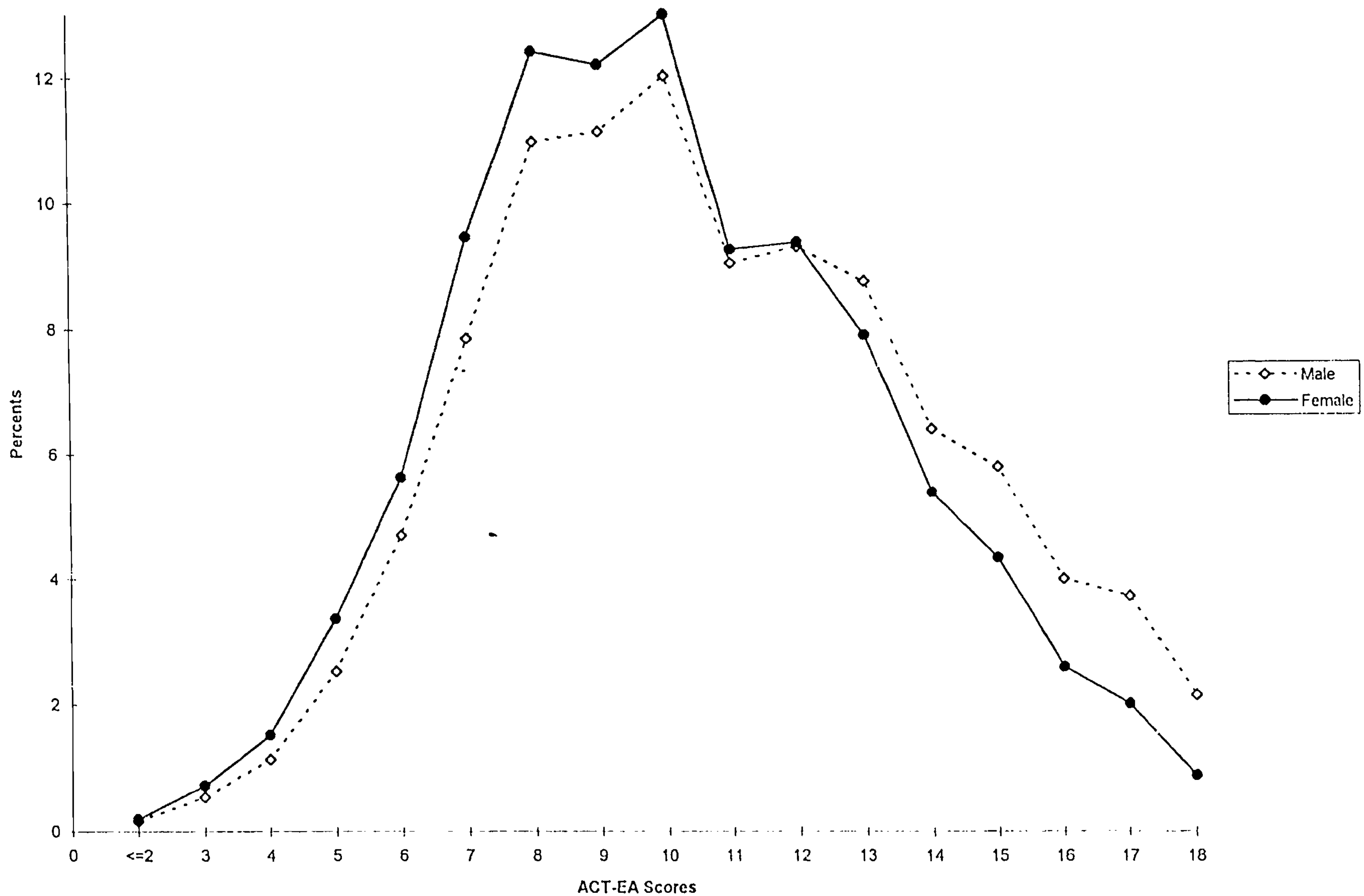


Figure 5. 1992 ACT-AG Score Distributions: Males' vs. Females'

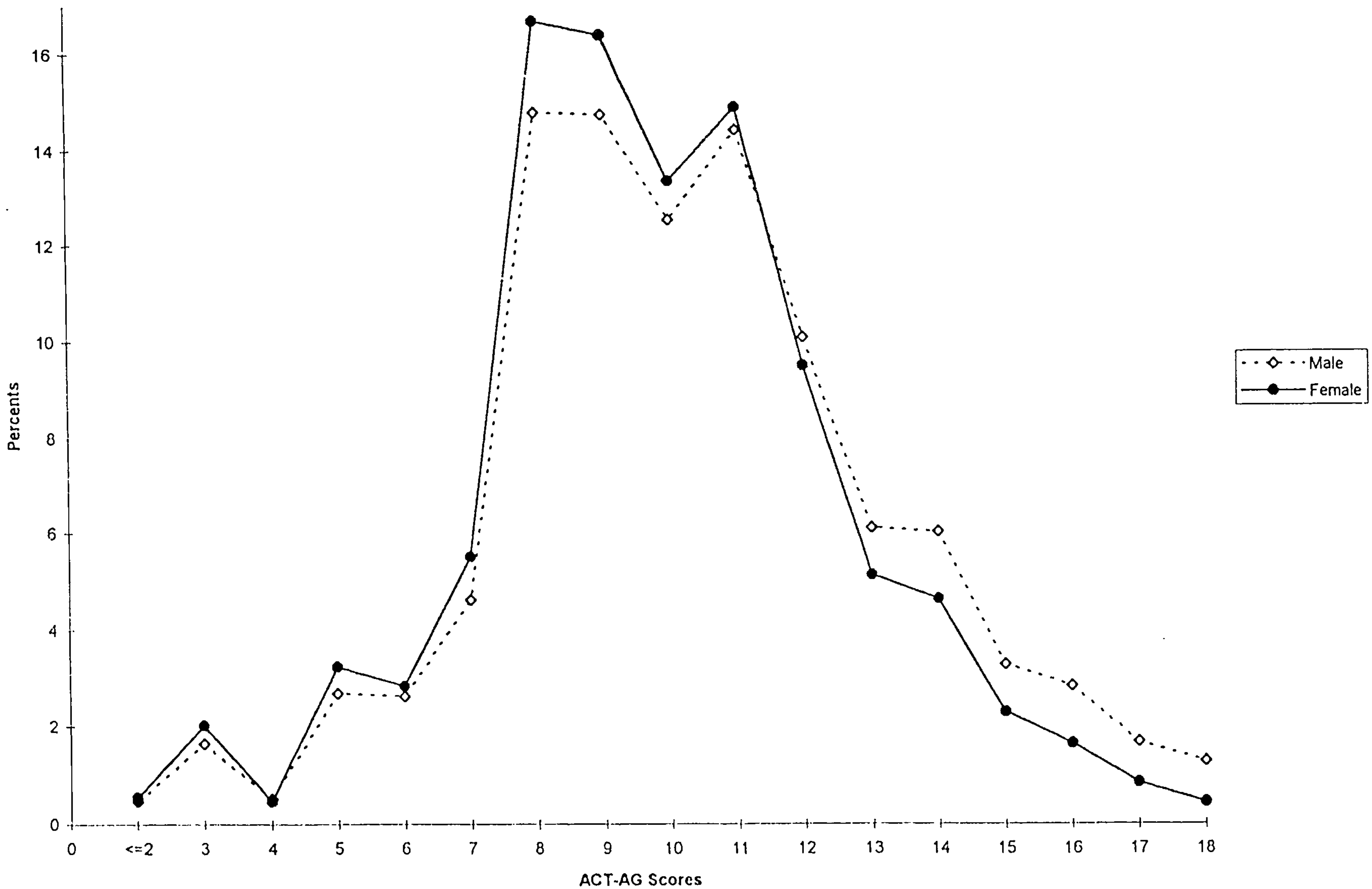


Figure 6. 1993 ACT-AG Score Distributions: Males' vs. Females'

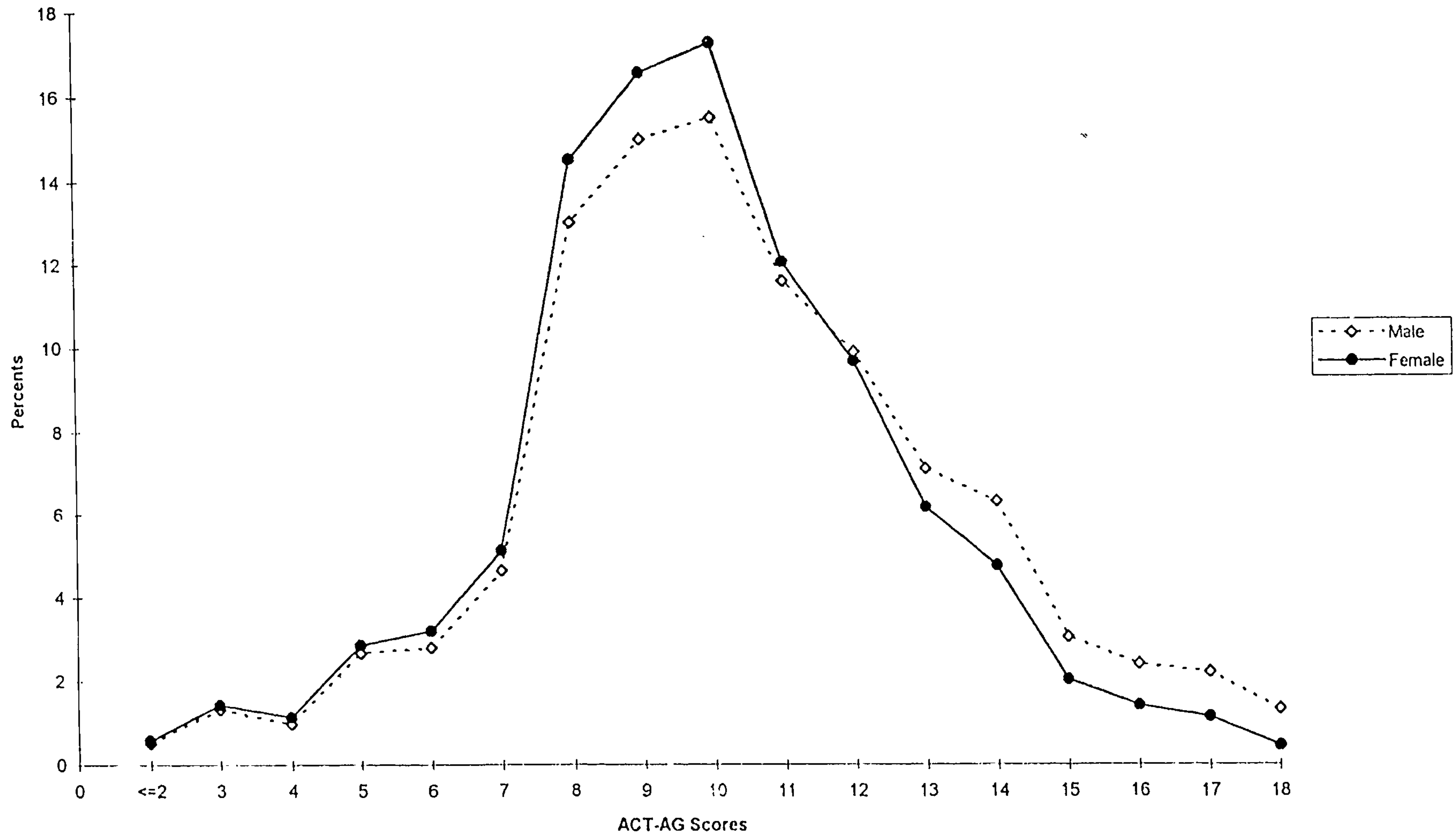


Figure 7. 1992 ACT-GT Score Distributions: Males' vs. Females'

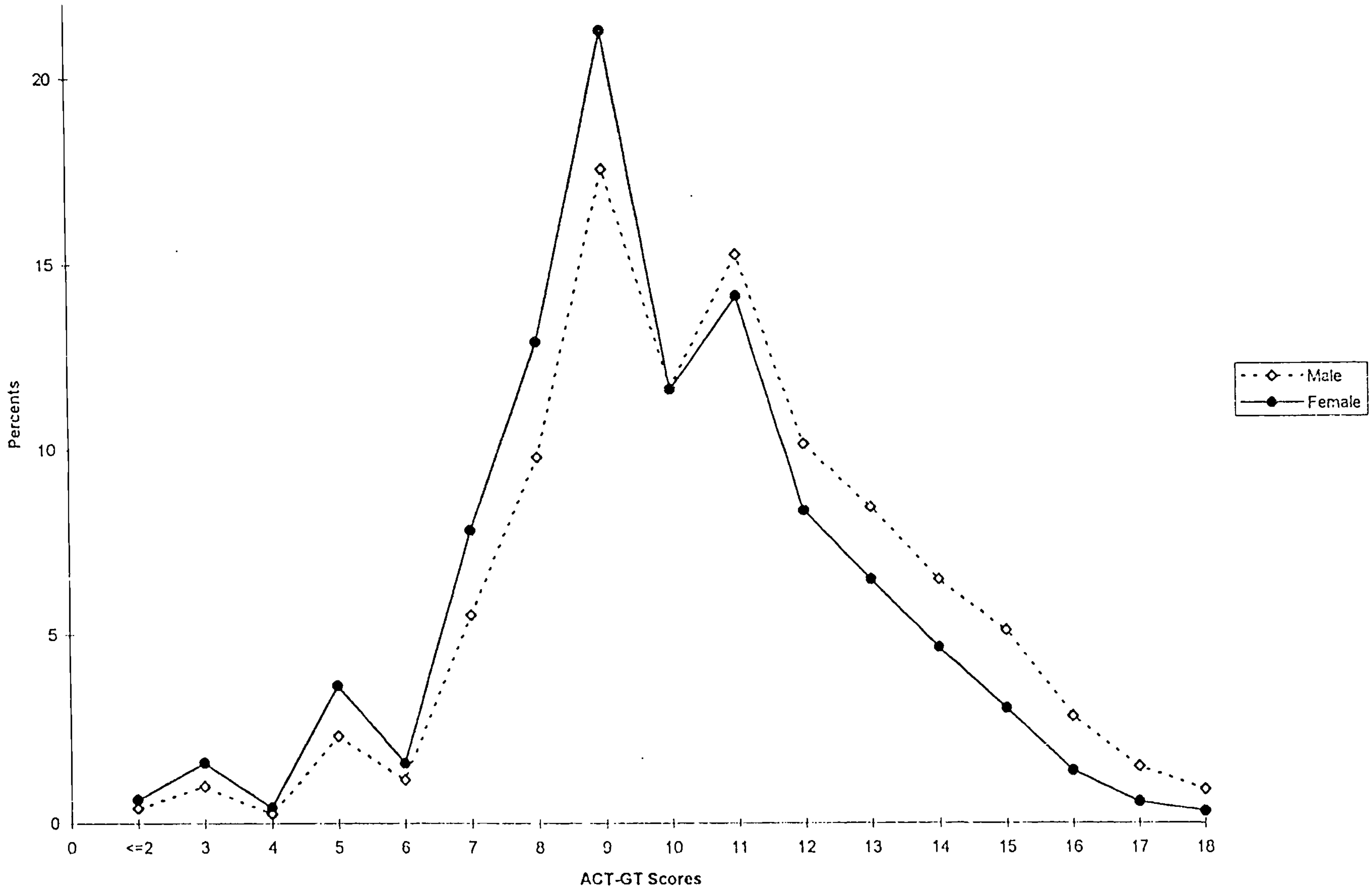


Figure 8. 1993 ACT-GT Score Distributions: Males' vs. Females'

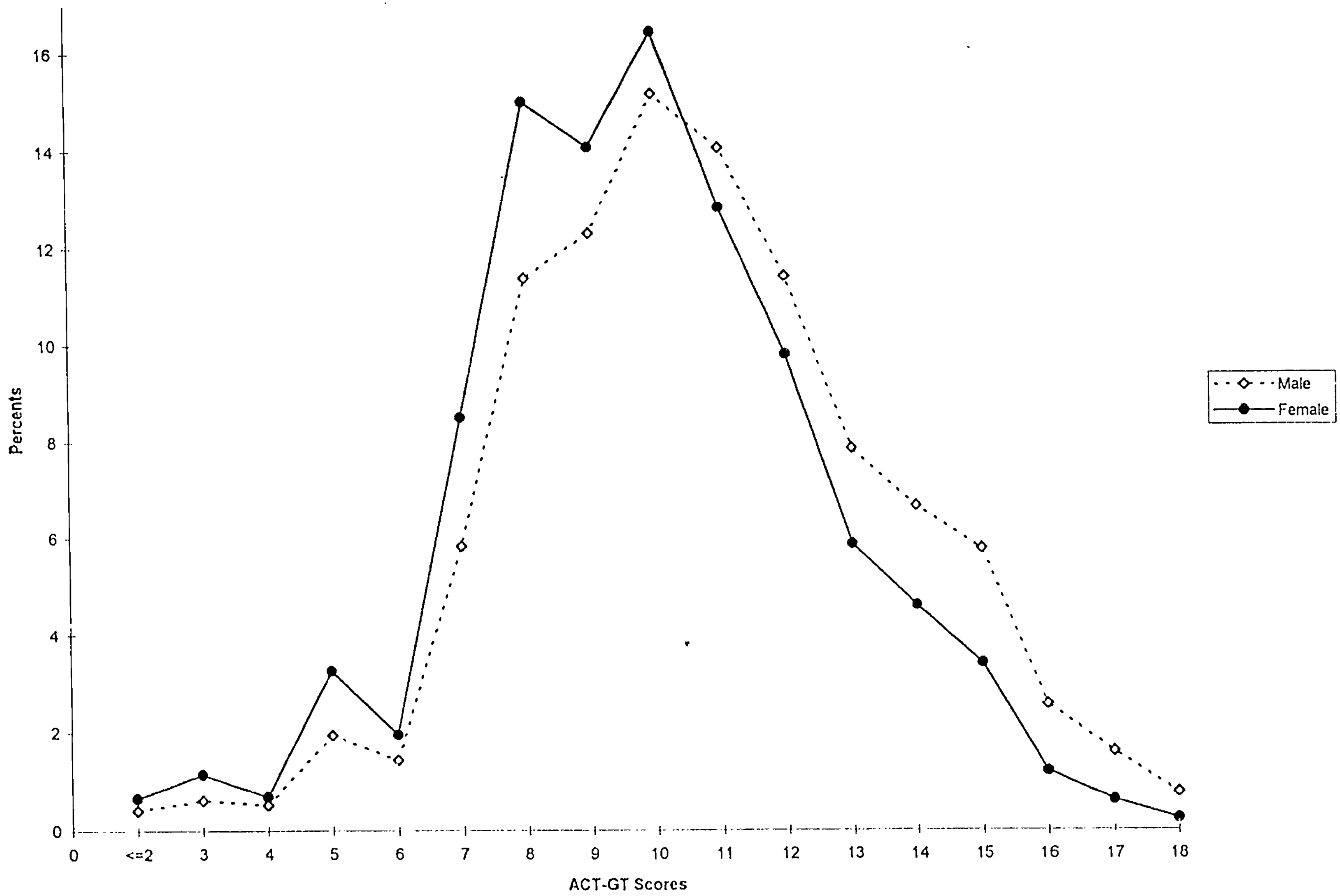


Figure 9. ACTM Score Distributions for Chinese Male and Female Students