DOCUMENT RESUME

ED 078 775 HE 004 371

AUTHOR	Doerann, Judith A.
TITLE	Equivalence of Grades in Multisection Courses.
INSTITUTION	Indiana Univ., Bloomington. Bureau of Educational Studies and Testing.
PUB DATE	73
NOTE	27p.; Indiana Studies in Prediction No. 22
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Academic Achievement; College Teachers; Comparative Statistics; *Grade Equivalent Scores; *Higher Education; *Performance; *Student Evaluation

ABSTRACT

This report investigates the equivalence of grades in multisection courses. Ten courses at Indiana University were identified for this study from many undergraduate, multisection courses offered in the fall semester, 1971. An attempt was made to select courses from every school with several from a variety of departments in Arts and Sciences. For all students in each of the selected courses, grades for the course and the Scholastic Aptitude Test (SAT) totals were obtained. The mean grade and mean SAT score for each course were used as the bases to study variations among related sections. Results indicated similarity in grading patterns across sections does not appear to be universally practiced. Appendices include statistical data. (MJM)

Equivalence of Grades in Multisection Courses

INDIANA STUDIES IN PREDICTION

NUMBER TWENTY-TWO

Bureau of Educational Studies and Testing
INDIANA UNIVERSITY / 1973

EQUIVALENCE OF GRADES IN MULTISECTION COURSES

by

Judith A. Doerann

INDIANA STUDIES IN PREDICTION NO. 22

Bureau of Educational Studies and Testing
Indiana University
Bloomington, Indiana
1973

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EQUIVALENCE OF GRADES IN MULTISECTION COURSES

The Problem

In large colleges and universities, it is a common practice to offer several sections in those courses which are in heavy demand. The various sections are often taught by different instructors. For this reason, the question of comparability of grades from section to section arises. If it can be assumed that high aptitude students tend to receive generally high grades, and that low aptitude students would generally receive low grades, to what extent does a given level of aptitude result in equivalent grades across sections of the same course? Would a student who received a grade of B in section 16 of a course have received the same grade in section 21? Are some instructors "easy graders" and some "hard graders"? Students sometimes contend that their grades depend as much on the instructor giving the grades as on their performance. Any objective evidence for or against that thesis would be of interest to faculty and students alike, and the purpose of this study was to obtain such evidence in the context of multisection courses.

Procedures

Ten courses were identified for this study from the many undergraduate, multisection courses offered in the fall semester of 1971. An attempt was made to select courses from every School, with several from a variety of departments in Arts and Sciences, the School offering a large number of multisection courses. Another consideration in selection was the number of sections per course. The courses identified had from three to twenty-seven sections each. There was one course each from the Schools of Music and Business, eight from Arts and Sciences, among which were one language course, two physical science courses, four humanities/social science courses, and one HPER course. The ten courses were coded A through J in this report of the study, approximately according to the number of sections per course.

For all students in each of the selected courses, grades for the course and Scholastic Aptitude Test (SAT) total scores were obtained. Previous studies have shown that SAT scores are positively related to success in work at Indiana University (Chase, et al., 1963). Grades were considered an achievement measure, and SAT scores an aptitude, or academic talent, measure. Grades were converted from letters to numerical equivalents (A = 4.0, B = 3.0, etc.). It should be noted that any grades that could not be converted to numerical equivalents were not functional for this study; students receiving P, W, WF, I, and S were therefore not included. A few

students did not have SAT total scores on record; they were also excluded.

Using numerical grade equivalents and SAT scores, several statistics were calculated for each course across all sections and for each section separately. These statistics were mean grade ($\bar{G}$) and grade standard deviation (S_G), mean SAT score ($\bar{A}$) and SAT standard deviation (S_A). The number (N) of students whose scores contributed to each of those statistics was obtained. For each course, the correlation (r_{GA}) between course grades and SAT scores was computed.

To determine whether any course in this study had undergone some systematic selection process to place students in particular sections, each department offering one of the courses was contacted. In most cases there were no prerequisites, and those which existed applied to all sections of a course. No assignment or selection procedures which would have resulted in systematic differences between sections were reported by the departments. Students who preregistered or enrolled early in the registration process had a better opportunity to select preferred section times, but it would be difficult to verify any effect of this on section characteristics. Consequently, for the purpose of this study, differences between sections of such factors as ability, interest, and motivation will be considered due to chance variations.

To facilitate discussion of differences in grades between sections of the same course, it was first necessary to control for aptitude differences between sections. This was made possible by using an equation derived by Frederick Davis, in his book Educational Measurements and Their Interpretation (1964). The numerical equivalents of grades in each section were converted to standard scores, with standard deviation equal to that of grades in all sections of a course combined [term (1) in Formula 1]. They are distributed around the mean of grades across all sections [term (2) in Formula 1]. Then, because grades are affected to some extent by differences in aptitude level, the scores were adjusted by a term reflecting the effect of the aptitude level of the section [term (3) in Formula 1]. The strength of the relationship between aptitude and achievement for a given course is expressed in the correlation coefficient (r_{GA}).

The equation used to control for aptitude differences so that grades could be compared across sections was

$$\text{adjusted grade equivalent} = \underset{(1)}{S_{Gc}\left(\frac{G_s-\overline{G}_s}{S_{Gs}}\right)} + \underset{(2)}{\overline{G}_c} + \underset{(3)}{S_{Gc}\, r_{GA}\left(\frac{\overline{A}_s-\overline{A}_c}{S_{Ac}}\right)} \qquad \text{[Formula 1]}$$

where $\overline{G}_c$ and S_{Gc} are the mean and standard deviation of grades across all sections of a course.

$\overline{G}_s$ and S_{Gs} are the mean and standard deviation of grades in a certain section.

G_s is a certain grade in a section; in this study, comparisons to grades of 2 (=C) and 3 (=B) were determined.

r_{GA} is the correlation between grades and SAT scores across all sections of a course.

$\bar{A}_c$ and S_{Ac} are the mean and standard deviation of SAT scores across all sections of a course.

$\bar{A}_s$ is the mean SAT score in a certain section.

Thus, for example, if a grade mean for a course is 2.8 and the mean is 3.5 for a section of that course, but the aptitude of students in the section is about average for the course, then it appears that it was easier to get an A in the section than in the course as a whole--the section instructor was an "easy grader," and an A in the section is "worth less" than most A's in the course as a whole. In this type of situation, the adjusted grade equivalent of A (=4) for the course given in this section might be 3.7. Similarly a B (=3) in the section might be comparable to 2.6.

Using Davis's equation, adjusted grade equivalents for each section for B = 3 and C = 2 were obtained. A few examples may illustrate the procedure.

In course A, section 2, $\bar{G}_c = 2.23$, $S_{Gc} = 1.14$, $\bar{G}_s = 2.33$, $S_{Gs} = 1.12$, $r_{GA} = 0.48$, $\bar{A}_c = 996.2$, $S_{Ac} = 172.1$, $\bar{A}_s = 1005.4$.

For 2.0:

$$\text{adjusted grade equivalent} = 1.14\left(\frac{2.0-2.33}{1.12}\right) + 2.23 + 1.14(0.48)\left(\frac{1005.4-996.2}{172.1}\right) = 1.92$$

In course C, section 3, $\bar{G}_c = 2.54$, $S_{Gc} = 0.96$, $\bar{G}_s = 2.15$, $S_{Gs} = 1.04$, $r_{GA} = 0.42$, $\bar{A}_c = 967.4$, $S_{Ac} = 172.0$, $\bar{A}_s = 975.2$.

For 3.0:

$$\text{adjusted grade equivalent} = 0.96\left(\frac{3.0-2.15}{1.04}\right) + 2.54 + 0.96(0.42)\left(\frac{975.2-967.4}{172.0}\right) = 3.34$$

Since the purpose of this study was to determine differences between sections of the same course, the mean grade and mean SAT score for each course were used as the bases to study variations among related sections. Variations found were attributed to differences in grading practices among instructors of the same course, and no attempt was made to compare courses or departments.

Results

The ten courses varied from three sections to twenty-seven sections per course. In addition to the variability in number of sections per course, the number of students per section varied from course to course and also from section to section in the same course. The following table lists courses coded from A to J, total number of students per course, number of sections, and mean number of students per section. The range of section size from lowest to highest is included to indicate variability within courses.

Table 1. Course and Section Enrollment Size

Course Code	Total Number of Students per Course	Number of Sections per Course	Number of Students in Section		
			Mean per Section	Lowest	Highest
A	544	3	181.3	130	214
B	306	3	102.0	21	164
C	462	3	154.0	20	221
D	238	5	47.6	31	62
E	1395	6	232.5	7	311
F	2044	11	185.8	44	234
G	224	13	17.2	14	21
H	454	13	34.9	22	50
I	362	17	21.3	17	26
J	341	27	12.6	4	21

From Table 1 it can be seen that average section size varies from 12.6 students in course J to 232.5 in course E. The number of sections made available and the consequent number of students per section are the result of departmental decisions. Anticipated course enrollment may be too large for effective instruction in one section, or instructional method may require a certain group size, such as a seminar. Those requirements would result in sections of a certain approximate size, but not differing

systematically from one another in other characteristics. A third possible reason for sectioning courses would be the desire to provide groups based on ability or interest, but none of the departments sponsoring courses in this study reported that they were operating on that principle.

The courses varied somewhat in average aptitude, indicated by mean SAT scores, and average grades. Table 2 summarizes these statistics and includes the correlation coefficient which indicates the strength of the relationship between grades, a measure of achievement, and SAT scores, a measure of aptitude. The range of mean grades of sections, shown in Table 2, is an indication of grade variability among sections within courses. Complete statistics for each course and for every section can be found in the Appendix.

Table 2. Mean SAT, Mean Grade, and Correlation by Course; Range of Section Mean Grades

Course	Course Statistics Across All Sections			Range of Mean Grades of Sections
	Mean SAT	Mean Grade	Correlation SAT/Grades	
A	996.2	2.23	0.48	2.05--2.33
B	1111.0	2.56	0.31	2.48--2.95
C	967.4	2.54	0.42	2.15--2.69
D	991.1	2.60	0.51	2.30--2.90
E	961.5	2.78	0.36	2.57--2.99
F	1014.7	2.44	0.50	2.28--2.69
G	914.3	3.51	0.08	3.00--4.00
H	1011.5	2.68	0.33	2.46--2.92
I	905.8	2.86	0.36	2.35--3.71
J	976.0	2.49	0.31	1.79--3.44

It is interesting to note that the mean grade for course G is considerably higher than the others, and that the grade/SAT correlation is very low. The low correlation could indicate that the SAT is a poor measure of aptitude for the subject taught in course G, or that there is no relation between a student's ability and his eventual grade. However, neither the high course mean, the low correlation, or other variations interfere with comparing grades

between sections of the same course; the course means are the standards for comparing section grades, and the correlation allows some correction for the effect, if any, of aptitude level differences. Therefore, differences in mean grades between courses do not affect the central findings in this study.

If sections within a course were graded uniformly, i.e., if there were no "easy" or "hard" graders, adjusted grade equivalents obtained with the equation would have been the same across all sections of a course. Sections graded more or less like the course as a whole would have adjusted grade equivalents close to 2.0 and 3.0, the two grades for which equivalents were computed in this study.

In the Appendix, adjusted grade equivalents for 2.0 and 3.0 are shown for each section in the last two columns. Table 3 summarizes some of that information; it shows for each course the number of sections whose adjusted grade equivalents for a grade of C, and for B, fell within given ranges. From this it can be seen that some courses have adjusted grade equivalents for sections that cluster rather closely to the standards of 2.0 and 3.0, while adjusted grades in others vary from course standards quite widely.

Table 3. Adjusted Grade Equivalents; Frequency Distribution of Sections

For Grade of 2.0 (C)

Interval of Grades Equivalent to 2.0 (C)	Courses									
	A	B	C	D	E	F	G	H	I	J
-1.0 to -.49							1			1
-0.5 to -.01										
0.0 to .49										
0.5 to .99							1		1	1
1.0 to 1.49							2		3	5
1.5 to 1.99	2	2	1	3	4	4	2	5	6	7
2.0 to 2.49	1	1	2	2	2	7	4	7	5	8
2.5 to 2.99							2	1	2	5

For Grade of 3.0 (B)

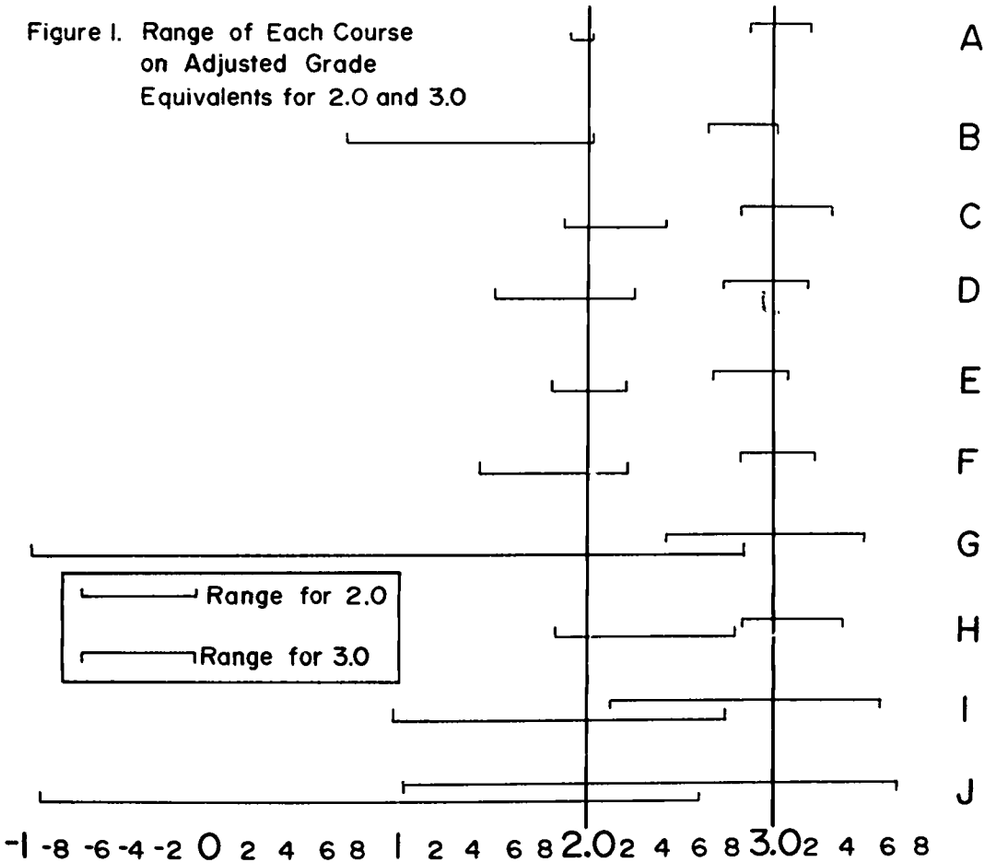
Interval of Grades Equivalent to 3.0 (B)	Courses									
	A	B	C	D	E	F	G	H	I	J
1.0 to 1.49							1			1
1.5 to 1.99										
2.0 to 2.49							1		2	1
2.5 to 2.99	2	2	1	1	4	6	3	7	8	11
3.0 to 3.49	1	1	1	4	2	5	7	6	5	11
3.5 to 3.99			1						2	3
4.0										

It should be remembered that adjusted grade equivalents have been corrected for any existing differences among sections in aptitude level, as measured by the SAT, and also for differences in grade mean between courses. The variation of adjusted grade equivalents indicates variation among course instructors in grading standards.

Figure 1 expresses the grading variation graphically. For each course, the lowest and highest adjusted grade equivalents obtained from any section in the course were plotted; this was done for equivalents of 2.0 and 3.0. Thus the figure indicates the ranges of adjusted grade equivalents of 2.0 and 3.0 for each course.

Uniformity in grading practice was evidently not the case in the courses selected, since there is a range of adjusted grade equivalents for each course. In some cases the range is quite extreme, particularly in courses G, I, and J. In a few cases the ranges of adjusted grade equivalents among sections for grades of 2.0 and 3.0 overlap; this means that at least one section of the course has an adjusted grade equivalent for 2.0 that is higher than the adjusted grade equivalent for 3.0 from another section in the same course. From this it can be inferred that it was more difficult to earn a C (=2.0) in the first section than to earn a B (=3.0) in the second. (Specific data for all sections are in the Appendix.) Course I, for example, has five sections with adjusted grade equivalents for 2.0 that are higher than the

Figure I. Range of Each Course on Adjusted Grade Equivalents for 2.0 and 3.0



lowest section equivalent for 3.0. In sections 1, 3, 10, 14, and 17 of course I it was apparently more difficult to earn a C than to earn a B in section 2. Another interpretation is that the section 2 instructor was an easier grader than the instructors of sections 1, 3, 10, 14, and 17.

In the Appendix table, numbers of sections that seem to have been generously graded are indicated by a single asterisk; sections taught by a "hard grader" are indicated by a double asterisk. There are a few instances where severity of grading varied within a section; for example, in course D, section 1, and course J, section 3, a C was relatively easy to earn, though a B was more difficult to obtain than in the course as a whole; in both of these sections the adjusted grade equivalents for 2.0 were less than 2.0 and for 3.0 were greater than 3.0.

Conclusions

Multisection courses vary in number of sections per course and number of students per section. Although departments request numbers of sections and thereby determine section size, none of the departments sampled placed students in certain sections. There was no obvious basis for sections to differ systematically from one another in grading standards.

However, sections did vary somewhat in ability as measured by the SAT. There was also considerable variation in grades from

section to section. Even when correction had been made for differences between sections in aptitude level, grades were not equivalent in general across most sections of the same course. Adjusted grade equivalents indicated that students would have as much difficulty earning a C in some sections as students with equal ability in other sections would have earning a B. Results from many sections showed adjusted grade equivalents relatively close to standards; in these sections grading was like that for the course as a whole. Nevertheless, similarity in grading patterns across sections does not appear to be universally practiced.

APPENDIX

Appendix.

	N	$\overline{G}$	S_G	$\overline{A}$	S_A		
Course A	544	2.23	1.14	996.2	172.1	r_{GA} = 0.48	
						Equivalents	
						2.0	3.0
Section 1	200	2.26	1.18	986.7	173.9	1.95	2.91
2	214	2.33	1.12	1005.4	169.2	1.92	2.89
3	130	2.05	1.11	975.4	174.7	2.18	3.20
Course B	306	2.56	0.81	1111.0	149.2	r_{GA} = 0.31	
						Equivalents	
						2.0	3.0
Section 1*	21	2.95	0.86	1138.4	166.9	1.71	2.65
2	121	2.59	0.77	1089.0	150.4	1.90	2.95
3	164	2.48	0.82	1123.7	144.8	2.11	3.10
Course C	462	2.54	0.96	967.4	172.0	r_{GA} = 0.42	
						Equivalents	
						2.0	3.0
Section 1	221	2.69	0.96	958.2	166.5	1.83	2.83
2	221	2.43	0.93	975.8	177.8	2.12	3.15
3**	20	2.15	1.04	975.2	171.5	2.42	3.34

Appendix, continued

	N	$\bar{G}$	S_G	$\bar{A}$	S_A		
Course D	238	2.60	0.97	991.1	171.3	r_{GA} = 0.51	
						Equivalents	
						2.0	3.0
Section 1	35	2.60	0.81	992.9	158.2	1.89	3.08
2	56	2.61	1.14	1015.9	185.5	2.15	3.00
3*	62	2.90	0.82	997.3	176.1	1.55	2.74
4	54	2.30	1.04	971.3	167.2	2.26	3.20
5	31	2.55	0.89	966.5	158.9	1.93	3.02
Course E	1395	2.78	0.81	961.5	166.4	r_{GA} = 0.36	
						Equivalents	
						2.0	3.0
Section 1	263	2.72	0.72	952.6	164.9	1.95	3.08
2	311	2.99	0.	947.7	171.0	1.81	2.77
3	278	2.71	C	975.6	157.8	2.10	3.09
4	225	2.91	0.78	959.4	175.6	1.83	2.87
5	311	2.90	0.82	973.9	163.4	1.91	2.90
6	7	2.57	1.13	871.9	126.6	2.21	2.93

Appendix, continued

	N	$\bar{G}$	S_G	$\bar{A}$	S_A		
Course F	2044	2.44	1.06	1014.7	165.8	r_{GA} = 0.50	
						Equivalents	
						2.0	3.0
Section 1	225	2.53	1.05	1022.4	156.2	1.93	2.94
2	234	2.31	1.17	1010.8	155.7	2.15	3.05
3	223	2.48	0.96	992.2	164.6	1.84	2.94
4	217	2.69	0.82	1033.6	165.3	1.61	2.90
5	191	2.29	0.98	100.8	170.8	2.08	3.16
6	165	2.64	0.88	999.6	179.4	1.62	2.81
7	205	2.28	1.26	1020.9	172.0	2.22	3.07
8	220.	2.31	0.97	1022.8	159.3	2.13	3.22
9	231	2.49	1.18	1024.4	169.9	2.03	2.93
10	89	2.46	1.19	1012.8	167.8	2.02	2.91
11	44	2.31	1.22	1019.5	171.4	2.19	3.05

Appendix, continued

	N	$\bar{G}$	S_G	$\bar{A}$	S_A		
Course G	224	3.51	0.83	914.3	138.2	r_{GA} = 0.08	
						Equivalents	
						2.0	3.0
Section 1*	16	3.75	0.58	927.2	113.0	1.01	2.44
2	20	3.55	0.94	895.6	137.2	2.13	3.02
3**	14	3.00	1.30	872.4	80.2	2.85	3.49
4**	17	3.12	0.78	946.2	156.6	2.33	3.40
5	18	3.44	0.70	948.6	114.1	1.82	3.00
6**	18	3.28	1.07	899.6	151.1	2.51	3.29
7	21	3.48	0.75	964.1	170.9	1.90	3.00
8*	20	3.60	0.50	929.5	119.7	0.86	2.52
9	17	4.00	0.00	932.1	162.4	--	--
10	16	3.56	0.63	881.2	140.6	1.44	2.76
11	17	3.71	0.99	889.8	165.2	2.06	2.90
12**	15	3.27	1.03	891.3	138.6	2.48	3.28
13*	15	3.87	0.35	881.1	104.2	-0.94	1.43

Note: No equivalents can be calculated for section 9 because there was no variation; all students received an A in that section. The lack of variation makes Davis's equation unusable there.

Appendix, continued

	N	$\bar{G}$	S_G	$\bar{A}$	S_A		
Course H	454	2.68	0.76	1011.5	153.8	r_{GA} = 0.33	
						Equivalents	
						2.0	3.0
Section 1	33	2.73	0.76	1022.5	140.5	1.97	2.97
2	50	2.82	0.77	996.8	154.7	1.85	2.83
3	38	2.92	0.78	1042.1	154.5	1.83	2.81
4	22	2.54	0.74	1024.4	164.7	2.15	3.17
5	25	2.52	0.87	985.8	170.5	2.18	3.07
6	47	2.55	0.65	1045.6	152.0	2.09	3.26
7	23	2.82	0.65	1059.4	145.6	1.80	2.97
8	22	2.86	0.94	1049.9	132.0	2.05	2.86
9	48	2.48	0.80	976.1	179.4	2.17	3.12
10	50	2.64	0.80	1004.8	153.7	2.06	3.01
11	39	2.74	0.75	986.9	154.2	1.89	2.90
12	29	2.72	0.70	995.3	142.9	1.87	2.96
13	28	2.46	0.58	992.4	121.8	2.05	3.36

Appendix, continued

	N	$\bar{G}$	S_G	$\bar{A}$	S_A		
Course I	362	2.86	0.97	905.8	173.7	r_{GA} = 0.36	
						Equivalents	
						2.0	3.0
Section 1**	21	2.48	1.03	889.8	180.8	2.38	3.33
2*	24	3.71	0.86	934.9	140.5	0.98	2.12
3**	24	2.42	1.18	943.5	195.4	2.59	3.41
4*	21	3.52	0.98	901.6	188.0	1.35	2.34
5	21	2.95	1.02	874.4	150.2	1.89	2.84
6	20	2.85	1.04	878.1	182.3	2.01	2.94
7	22	2.91	0.81	905.7	157.9	1.77	2.97
8	20	2.95	0.89	839.4	135.6	1.69	2.78
9	21	2.81	0.75	859.7	152.1	1.72	3.01
10**	22	2.41	1.18	996.0	179.1	2.70	3.53
11	24	2.96	0.62	886.5	167.6	1.32	2.88
12	23	2.61	0.78	865.7	201.2	2.02	3.26
13	26	3.08	0.89	926.9	161.8	1.73	2.82
14**	21	2.48	0.81	897.6	195.3	2.27	3.49
15	18	3.11	0.68	979.8	172.9	1.43	2.85
16	17	2.82	0.88	940.6	209.4	1.89	2.99
17**	17	2.35	0.79	872.2	144.8	2.36	3.59

Appendix, continued

	N	$\overline{G}$	S_G	$\overline{A}$	S_A		
Course J	341	2.49	1.37	976.0	164.7	r_{GA} = 0.31	
						Equivalents	
						2.0	3.0
Section 1	14	2.64	1.34	1063.6	157.0	2.06	3.08
2	19	2.37	1.38	947.1	189.8	2.05	3.04
3	10	2.70	1.16	1061.0	119.3	1.88	3.06
4	14	2.43	1.45	930.1	139.9	1.97	2.91
5*	11	3.18	1.17	1103.8	169.9	1.44	2.61
6**	10	1.80	1.48	949.3	183.2	2.61	3.53
7*	13	2.92	1.32	948.5	118.9	1.46	2.50
8**	13	2.00	1.22	965.8	169.2	2.43	3.69
9*	14	2.79	1.25	921.4	201.5	1.48	2.58
10	13	2.24	1.54	952.8	163.4	2.22	3.11
11	18	2.67	1.33	989.2	137.3	1.83	2.86
12**	10	1.90	1.66	951.7	168.0	2.51	3.34
13**	12	2.17	1.59	1050.2	183.8	2.53	3.40
14	21	2.38	1.32	954.3	204.7	2.24	3.08
15*	9	3.44	0.73	724.1	109.0	-0.86	1.02
16*	13	2.85	1.21	969.7	187.2	1.51	2.64
17**	16	1.94	1.61	1002.1	155.0	2.61	3.46
18	12	2.92	1.51	1066.6	107.0	1.89	2.80
19	10	2.30	1.34	996.2	115.1	2.24	3.26

Appendix, continued

	N	$\overline{G}$	S_G	$\overline{A}$	S_A	Equivalents	
Course J, continued						2.0	3.0
Section 20	12	2.33	1.44	971.8	197.2	2.17	3.12
21	12	2.67	0.87	954.4	119.6	1.40	2.94
22*	9	3.11	1.05	993.3	133.6	1.09	2.39
23	15	2.60	1.50	982.2	139.8	1.96	2.87
24	13	2.85	1.34	1023.3	154.1	1.74	2.77
25*	4	3.00	0.82	1022.5	115.3	0.94	2.61
26**	10	2.00	1.33	994.2	123.1	2.54	3.57
27**	14	1.79	1.63	892.8	111.3	2.45	3.29