DOCUMENT RESUME

ED 348 369 TM 014 328

AUTHOR Pike, Gary R.
TITLE A Comparison of the College Outcome Measures Program (COMP) Exam and the ETS Academic Profile.
INSTITUTION Tennessee Univ., Knoxville. Learning Research Center.
PUB DATE Aug 88
NOTE 18p.; This document is Appendix B of the 1987-88 Performance Funding Report for the University of Tennessee, Knoxville, and is included as pages 63-79 of that document.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Ability; *Achievement Tests; *College Seniors; *Comparative Testing; Construct Validity; Educational Assessment; Higher Education; Outcomes of Education; *Test Format; Test Use
IDENTIFIERS *Academic Profile; *College Outcome Measures Project; Test Appropriateness; University of Tennessee Knoxville

ABSTRACT

Two of the most popular assessment instruments for measuring college student educational outcomes are the College Outcome Measures Program (COMP) of the American College Testing Program (1987) and the Academic Profile of the Educational Testing Service (ETS). Characteristics of these tests were compared, and the sensitivity of each test to students' educational experiences was evaluated. In the fall and winter quarters of the 1987-88 school year, seniors at the University of Tennessee (Knoxville) took one of these tests. Overall, more seniors took the COMP (n=1,261) than the Academic Profile (n=810). Thirty-eight seniors took both examinations. The results indicate that both tests are unidimensional measures that are not particularly sensitive to students' general education experiences. Both tests represented measures of individual differences and were powerful measures of academic ability. Because of their sensitivity to individual differences, both the COMP and the Academic Profile are not considered appropriate for evaluating the impact of general education programs. Institutions would not be well-advised to use the subscales of either examination to evaluate the components of a curriculum. Six tables give study results. (SLD)

A Comparison of the College Outcome Measures Program (COMP) Exam And the ETS Academic Profile

Gary R. Pike

APPENDIX B OF THE

1987-88

PERFORMANCE FUNDING REPORT

FOR

THE UNIVERSITY OF TENNESSEE, KNOXVILLE

Compiled By

Trudy W. Banta

Research Professor

With

Technical Reports

By

Gary R. Pike
Associate Director
Assessment Resource Center

LEARNING RESEARCH CENTER
1819 Andy Holt Avenue
Knoxville, Tennessee
37996-4350

August 1988

A COMPARISON OF THE COLLEGE OUTCOME MEASURES PROGRAM (COMP) EXAM AND THE ETS ACADEMIC PROFILE

Gary R. Pike
Associate Director
Assessment Resource Center

During the 1980s, the reports of national advisory commissions, coupled with the actions of state higher education commissions and regional accrediting associations, have sparked intense interest in the assessment of student educational outcomes (El-Khawas, 1987; Ewell, 1987). Responding to the heightened interest in student learning, test developers have introduced several new outcomes measures, many designed to assess the effectiveness of general education programs (Pike, in press).

Because so many assessment instruments are available, it is critical that institutions utilize outcomes measures that accurately reflect their curricula, missions, and student characteristics. Judging the appropriateness of an assessment instrument requires that the content, psychometric properties, and construct validity of the instrument be carefully examined.

Two of the most popular assessment instruments are the College Outcome Measures Program examination (American College Testing Program, 1987) and the ETS Academic Profile (ETS College and University Programs, 1988). Both the COMP exam and the Academic Profile are designed to measure the outcomes of general education. Because it is difficult, if not impossible, to identify a core of knowledge (content) that is common to general education programs at most colleges and universities, both the COMP exam and the Academic Profile minimize the need to recall specific facts. However, according to staff at ACT and ETS, familiarity with content improves test performance.

The objective form of the COMP exam takes approximately 2 1/2 hours to administer and contains 60 questions, each with two correct answers. These questions are divided among 15 separately timed activities drawing on materials (stimuli) from television programs, radio broadcasts, and print media. Students taking the COMP exam are instructed that there is a penalty for guessing (i.e., incorrect answers will be subtracted from st dents' scores). The combination of two correct answers for each question and the guessing penalty means that each question on the COMP exam is worth 4 points. The maximum possible score on the COMP exam is 240 points.

In addition to a total score, the COMP exam provides three content subscores (Functioning within Social Institutions [FSI], Using Science and Technology [UST], and Using the Arts [UA]) and three process subscores (Communicating [COM], Solving Problems [SP], and Clarifying Values [CV]). It is difficult to determine precisely what underlying constructs these scales are designed to measure because the technical manual for the COMP exam only provides one-paragraph descriptions of the subscales (Forrest & Steele, 1982).

The three-hour version of the Academic Profile consists of 144 questions and is designed to provide scores for individuals. The maximum possible score on the Academic Profile is 144 points. The questions on the Academic Profile are designed to measure four skills (Reading [READ], Writing [WRIT], Critical

Thinking [CT], and Mathematics [MATH]) across three content areas (Humanities [HUM], Social Sciences [SS], and Natural Sciences [NS]). As is the case with the COMP exam, ETS provides little information about the constructs the seven subscales of the Academic Profile are designed to measure. Students taking the Academic Profile are told that there is no penalty for guessing. Like the COMP exam, the Academic Profile uses multiple stimuli. However, all of the stimuli in the Academic Profile consist of written passages.

As part of its performance funding effort, the University of Tennessee - Knoxville has begun a study to compare the COMP exam and the Academic Profile. What follows is a report on these research efforts. This report is organized around two topics: (1) a comparison of the characteristics of each test, and (2) an evaluation of the sensitivity of each test to students' educational experiences. In reading this report, it is important to remember that these evaluations represent only one part of an ongoing effort to analyze a variety of assessment instruments.

Procedures

The data for this research were gathered during the Fall and Winter quarters of the 1987-88 academic year. During the Fall quarter, 1064 seniors were tested using either the objective form of the COMP exam or the three-hour version of the Academic Profile. Slightly more students took the COMP exam (598) than the Academic Profile (466). During the Winter quarter, 1007 students were tested. Again, more students took the COMP exam (663) than the Academic Profile (344).

Assignment to a testing group was based on two criteria: First, all students who had taken the COMP exam as freshmen were assigned to the COMP testing group as seniors. Second, students who had not been tested as freshmen were randomly assigned to either the COMP or the Academic Profile testing groups.

During the Winter quarter, 38 seniors agreed to take both the COMP exam and the Academic Profile. All of these students were volunteers, and they were compensated for their participation. Despite the fact that all of the students in this sample were self selected, their entering ability levels were close to the University average.

Comparing the characteristics of the COMP exam and the Academic Profile involved examining the psychometric properties of the tests, both independently and in combination. These analyses provided descriptions of student performance on the two tests and provided measures of the interrelationships among subscales. By examining these interrelationships, it is possible to draw inferences about what is being measured by the two tests.

In order to evaluate the sensitivity of the COMP exam and the Academic Profile to students' educational experiences, the scores of UTK students on the two tests were correlated with measures of course-taking. Coursework measures were derived from a series of supplemental items students completed when they took the COMP exam or the Academic Profile. Although the questions on coursework differed slightly depending on the test, the questions measured the amount of coursework in four areas: humanities, social sciences, mathematics,

and natural sciences. For students taking the Academic Profile, natural sciences coursework and mathematics coursework were combined

Test Characteristics

Score Profiles

The data presented in Table 1 describe the performance of UTK students on the COMP exam and the Academic Profile. Test data include mean scores ($\bar{X}$) and mean percentage correct scores ($\bar{X}$%) for each instrument over the two testing periods. In order to make scores on the two tests comparable, the mean percentage correct for norm groups ($\bar{NM}$%) and ratios of the UTK means to these norms ($\bar{X}$%/$\bar{NM}$%) also are provided. Standard deviations (Sx) are included as indicators of score dispersion.

Table 1:
Student Performance on the COMP Exam and the Academic Profile

COMP EXAM

	Fall (N=598)					Winter (N=663)				
	$\bar{X}$	$\bar{X}$%	$\bar{NM}$%	$\bar{X}$%/$\bar{NM}$%	Sx	$\bar{X}$	$\bar{X}$%	$\bar{NM}$%	$\bar{X}$%/$\bar{NM}$%	Sx
FSI	63.58	79	78	1.01	5.66	61.35	77	78	.99	6.68
UST	64.60	81	78	1.03	5.74	64.12	80	78	1.03	6.48
UA	62.59	78	75	1.04	6.32	60.12	75	75	1.00	5.37
COM	54.26	75	73	1.03	7.22	52.39	73	73	1.00	7.22
SP	77.92	85	78	1.09	5.85	75.44	79	78	1.01	6.66
CV	58.36	81	79	1.02	5.70	57.96	80	79	1.01	5.41
TOTAL	190.77	79	77	1.03	14.89	185.65	77	77	1.00	15.02

ACADEMIC PROFILE

	Fall (N=466)					Winter (N=344)				
	$\bar{X}$	$\bar{X}$%	$\bar{NM}$%	$\bar{X}$%/$\bar{NM}$%	Sx	$\bar{X}$	$\bar{X}$%	$\bar{NM}$%	$\bar{X}$%/$\bar{NM}$%	Sx
HUM	29.76	62	50	1.24	7.94	28.72	60	50	1.20	8.42
SS	28.89	60	50	1.20	7.79	28.04	58	50	1.16	7.95
NS	28.72	60	50	1.20	7.97	28.27	59	50	1.18	8.03
READ	23.03	64	50	1.28	6.25	22.20	62	50	1.24	6.61
WRIT	23.42	65	50	1.30	6.10	22.68	63	50	1.26	6.57
CT	19.43	54	50	1.08	6.49	18.86	52	50	1.04	6.37
MATH	21.50	60	50	1.20	6.20	21.28	59	50	1.18	6.19
TOTAL	87.38	61	50	1.22	22.13	85.02	59	50	1.18	22.81

Norms for the COMP exam were obtained from national statistics on the most recent administrations of the test (American College Testing Program, 1988). Norms for the Academic Profile were derived from ETS test specifications. Preliminary research by ETS (personal communication) indicates that these estimates may be slightly overstated, and the test may be more difficult than expected. This is particularly true for the Critical Thinking and Mathematics subscales.

An examination of the means and percent correct scores suggests that UTK students performed better on the COMP exam than on the Academic Profile. Total scores averaged 79 percent correct for the COMP exam in the Fall and 77 percent correct in the Winter. By way of comparison, total scores on the Academic Profile averaged 61 percent correct in the Fall and 59 percent correct in the Winter. On the six subscales of the COMP exam, Fall scores averaged between 75 percent correct (Communicating) and 85 percent correct (Solving Problems). Winter scores averaged between 73 percent correct (Communicating) and 80 percent correct (Clarifying Values and Using Science and Technology). For the seven subscales of the Academic Profile, Fall scores averaged between 54 percent (Critical Thinking) and 65 percent correct (Writing). For the Winter administration, scores averaged between 52 percent correct (Critical Thinking) and 63 percent correct (Writing).

A comparison of norms for the COMP exam and the Academic Profile refutes the suggestion that students perform better on the COMP exam. The norms presented in Table 1 indicate that the Academic Profile is much more difficult than the COMP exam. The mean percentage correct on the Academic Profile is assumed to be 50 percent, while the mean percentage correct reported for the COMP exam is 77 percent. Even greater disparities in difficulty levels can be found for the subscales of the two exams.

When the mean percentage correct scores for UTK students are expressed as ratios of national norms, it becomes obvious that student performance on the Academic Profile is superior to performance on the COMP exam. During the Fall quarter, the ratio of the UTK mean percentage correct to the national percentage correct was 1.22 for total scores on the Academic Profile and 1.03 for total scores on the COMP exam. For the Winter quarter, these ratios were 1.18 and 1.00 respectively.

Differences in student performance on the Academic Profile and the COMP exam cannot be attributed to differences in the entering ability levels of the two groups. The average ACT Assessment score for students taking the COMP exam was higher than the average for students taking the Academic Profile (22.04 and 21.47 respectively).

The data presented in Table 1 also suggest that there is greater variability in scores on the Academic Profile than on the COMP exam. For the Academic Profile, the standard deviations for total scores were 22.13 in the Fall quarter and 22.81 in the Winter quarter. For the COMP exam, the standard deviations for total scores were 14.89 and 15.02 respectively. Differences in variability are less dramatic for the subscales of the two exams, but they are present nonetheless.

The finding that students' scores on the Academic Profile were more variable than scores on the COMP exam is surprising given the fact that the theoretical range of the COMP exam (0 to 240) is much greater than the range of

the Academic Profile (0 to 144). Even if what are, in effect, negative scores on the COMP exam (0 to 120) are excluded, the observed differences in variability would be surprising.

One possible explanation for less variability in COMP exam scores is the low level of test difficulty for the COMP exam. The fact that the COMP exam is a relatively easy test creates a ceiling effect which restricts the range of variation in students' scores at the top of the scale. One practical result of the greater variability in scores on the Academic Profile is that the test is better able to differentiate among students than is the COMP exam.

The fact that the COMP exam is less difficult, and evidences less variability, than the Academic Profile has important practical consequences for the interpretation of test scores. Low difficulty levels, coupled with low levels of score variance and the fact that each question on the COMP exam can be worth as many as 4 points, create a situation in which small changes in student performance can have enormous effects on percentile ranks. For example, a change in students' responses on 1 of the 60 questions (2 responses) on the COMP exam would produce a change in the mean total score of 4 points (out of a possible 240 points). This 4 point score change is less than a 2 percent change in the possible score, but it translates into approximately a 10 percentile point gain or loss for scores between the 30th and 70th percentiles. The score decline of slightly over 5 points from the Fall to the Winter quarters for UTK seniors (see Table 1) translates into almost a 15 percentile point decline from the 60th to the 45th percentile.

Relationships Among Subscales

Examining the relationships among the subscales of an assessment instrument can provide useful information about whether that instrument actually measures the outcomes it purports to measure (Messick, 1987). For the purposes of the present investigation, the issue is whether the subscales of the COMP exam and the Academic Profile actually measure distinct, although possibly related, aspects of general education.

The first step in evaluating the relationships among subscales involved calculating correlations among subscales. Because both the COMP exam and the Academic Profile use the same questions to measure content and process/skill areas, correlations had to be calculated separately for the subscales within the content areas and within the process/skill areas. Correlations between content and process/skill subscales would not be statistically valid indicators.

Correlations among the subscales of the COMP exam and the Academic Profile are presented in Table 2. These correlations are based on data from both the Fall and Winter testing periods. Coefficients in the first subtable depict relationships among subscales of the COMP exam and are based on national data. This data was obtained from the ACT technical report on the COMP exam (Forrest & Steele, 1982). The coefficients in the first subtable are the same for both testing periods.

Table 2:
Correlations Among Subscores on the COMP Exam and the Academic Profile

COMP EXAM - NATIONAL SCORES

	Fall						Winter					
	FSI	UST	UA	COM	SP	CV	FSI	UST	UA	COM	SP	CV
FSI	.63	.59	.55				.63	.59	.55			
UST	.91	.68	.59				.91	.68	.59			
UA	.86	.88	.66				.86	.83	.66			
COM				.66	.56	.54				.66	.56	.54
SP				.84	.68	.57				.84	.68	.57
CV				.82	.85	.66				.82	.85	.66

COMP EXAM - UTK SCORES

	Fall						Winter					
	FSI	UST	UA	COM	SP	CV	FSI	UST	UA	COM	SP	CV
FSI	.61	.56	.42				.69	.54	.46			
UST	.98	.63	.41				.78	.70	.50			
UA	.66	.63	.67				.78	.85	.49			
COM				.75	.48	.42				.73	.48	.47
SP				.73	.58	.43				.70	.65	.50
CV				.60	.70	.65				.70	.79	.62

ACADEMIC PROFILE - UTK SCORES

	Fall							Winter						
	HUM	SS	NS	READ	WRIT	CT	MATH	HUM	SS	NS	READ	WRIT	CT	MATH
HUM	.84	.82	.79					.88	.85	.78				
SS	.98	.83	.82					.98	.86	.80				
NS	.94	.99	.84					.90	.93	.86				
READ				.81	.78	.82	.60				.83	.81	.81	.63
WRIT				.96	.80	.75	.64				.98	.83	.76	.60
CT				1.00	.93	.81	.67				.99	.93	.80	.67
MATH				.74	.80	.83	.80				.78	.74	.84	.79

Above Diagonal = Correlations
Diagonal = Reliabilities
Below Diagonal = Correlations (Disattenuated)

The second subtable presents data concerning the relationships among the subscales of the COMP exam, and the third subtable presents data concerning the relationships among the subscales of the Academic Profile. The data in the second and third subtables are based on the test scores of UTK seniors.

Coefficients above the diagonals in Table 2 are product-moment correlations. Coefficients on the diagonal of each matrix are reliability estimates for the subscales. Reliability estimates for the scores from the national sample were calculated using alpha reliability (Forrest & Steele, 1982). Reliability estimates for the scores of UTK students were calculated using a form of the KR-20 reliability coefficient (Gulliksen, 1950). These estimates are based on subscale means and variance statistics. The KR-20 reliability coefficient used in this research assumes homogeneity of item difficulty levels. The coefficients that are below the diagonals in Table 2 are correlations that have been corrected for attenuation (scale unreliability). These disattenuated correlations were calculated using the formula suggested by Gulliksen (1950).

The correlations presented in Table 2 provide an indication of the extent to which subscales represent unique aspects of the educational process. Highly intercorrelated subscales are undesirable because high correlations suggest that the subscales are measuring the same outcomes, rather than different aspects of general education.

Examination of the correlations based on the test scores of UTK students indicates that the subscales of the Academic Profile are more highly intercorrelated than the subscales of the COMP exam. Higher correlations among the subscales of the Academic Profile were found for both the Fall (.60 to .82 versus .41 to .56) and Winter quarters (.60 to .85 versus .46 to .54). Even when correlations based on national norms for the COMP exam (.54 to .59) are included in these comparisons, the disparity is not significantly reduced.

Even though the subscales of the Academic Profile are more highly intercorrelated than the subscales of the COMP exam, the correlations among the subscales of the COMP exam still are quite high. Moreover, disparities between intercorrelations for the two tests may be the result of higher levels of reliability for the Academic Profile.

An examination of the reliability coefficients on the diagonals of each matrix clearly shows that the subscales of the Academic Profile are more reliable than the subscales of the COMP exam. The range of KR-20 reliability estimates for the Academic Profile is .80 to .84 for the Fall quarter and .79 to .88 for the Winter quarter. The range of KR-20 reliability estimates for the COMP exam is .54 to .75 for the Fall and .49 to .73 for the Winter quarter. Even when national norms for the COMP exam are considered, the subscales of the Academic Profile are still more reliable.

Disattenuated correlations provide a means of compensating for differences in scale reliability. An examination of the correlations that have been corrected for attenuation reinforces the view that the subscales of both the Academic Profile and the COMP exam are highly correlated. During the Fall quarter, the range of disattenuated correlations for the Academic Profile was .74 to 1.00, while the range of disattenuated correlations for the COMP exam was .60 to .90 when the scores of UTK students were used. During the Winter

quarter, the range of disattenuated correlations was from .74 to .99 for the Academic Profile and from .70 to .85 for the COMP exam. When the results of national research were used, the range of disattenuated correlations was .82 to .91 for the COMP exam.

Because data were available for both the Fall and Winter quarters, it is possible to evaluate the stability of correlations and reliabilities for the COMP exam and the Academic Profile. Although the range of correlations among subscales was slightly lower for the Winter quarter, all of the subscales on the COMP exam were moderately intercorrelated for both quarters. Moreover, those scales with low correlations during the Fall quarter (e.g., Communicating and Clarifying Values) had low correlations in the Winter quarter. Similarly, those scales with higher correlations in the Fall quarter (Functioning within Social Institutions and Using Science and Technology) had high correlations in the Winter quarter.

While the correlations among subscales on the COMP exam were relatively stable, the reliability estimates for these subscales were much less stable. During the Fall quarter, the reliability estimate for Using the Arts was relatively high (.67). During the Winter quarter, this subscale evidenced the lowest level of reliability (.49). For both quarters, Communicating had the highest levels of reliability (.75 and .73 respectively).

For the Academic Profile, both the correlations among subscales and the reliability estimates for those subscales were remarkably stable across testing periods. For example, the content subscales of the Academic Profile were among the most highly correlated and the most reliable subscales irrespective of time period. On the other hand, the Mathematics subscale was one of the least reliable subscales, and it had relatively low correlations with the other subscales.

While a visual examination of the correlations among subscales suggests that both the COMP exam and the Academic Profile are measuring a single dimension, it is possible that other dimensions are present. In order to test whether there are unique aspects of the educational experience that are being measured by the COMP exam and the Academic Profile, principal components analyses were performed. Because of the moderate to high product-moment correlations among subscales, analyses were conducted only for those correlations. The presence of a single component in these analyses obviously would be replicated if disattenuated correlations were used.

In interpreting the results of these analyses, it was assumed that the existence of a single principal component would support the view that subscales measure a single outcome. The presence of more than one principal component would support the view that there are several outcomes being measured by the two tests.

Results of the principal components analyses for the content and process/skills areas of the COMP exam and Academic Profile are presented in the four subtables of Table 3. Results support the existence of a unidimensional structure. In no case does a meaningful second principal component emerge in the data analyses. Each subtable contains the factor loadings and communalities (estimates of explained variance) for each test subscore. Separate results are reported for the Fall and Winter quarters.

Table 3:
Principal Components Analysis of the Subscales for the COMP Exam and the Academic Profile

	COMP CONTENT AREAS			
	Fall		Winter	
	PATTERN LOADING	COMMU-NALITY	PATTERN LOADING	COMMU-NALITY
FSI	.83	.69	.81	.67
UST	.83	.69	.84	.70
UA	.74	.55	.80	.64

	COMP PROCESS AREAS			
	Fall		Winter	
	PATTERN LOADING	COMMU-NALITY	PATTERN LOADING	COMMU-NALITY
COM	.80	.65	.80	.64
SP	.81	.65	.81	.66
CV	.77	.59	.81	.66

	ACADEMIC PROFILE CONTENT AREAS			
	Fall		Winter	
	PATTERN LOADING	COMMU-NALITY	PATTERN LOADING	COMMU-NALITY
HUM	.93	.87	.94	.88
SS	.94	.89	.95	.89
NS	.93	.87	.92	.84

	ACADEMIC PROFILE PROCESS AREAS			
	Fall		Winter	
	PATTERN LOADING	COMMU-NALITY	PATTERN LOADING	COMMU-NALITY
READ	.91	.82	.92	.85
WRIT	.90	.81	.90	.80
CT	.92	.84	.92	.84
MATH	.81	.66	.81	.65

For the content subscales of the COMP exam, analysis of the data from the Fall quarter produced an eigenvalue of 1.93 for the first principal component. An eigenvalue of .63 was obtained for the second principal component. For the data from the Winter quarter, the eigenvalues of the first and second principal components were 2.00 and .54 respectively. An examination of the data presented in the first subtable indicates that all three content subscales had high loadings on the first principal component, and that component was able to explain a substantial proportion of the variance in each subscale.

The eigenvalues for the principal components underlying process subscores on the COMP exam also suggested a unidimensional structure. For the data from the Fall quarter, eigenvalues of 1.89 and .60 were obtained for the first and second principal components. For the data from the Winter quarter, the eigenvalues were 1.96 and .53 respectively. Here again, the process subscores had significant positive loadings on the first principal component, and this component was able to explain a substantial proportion of the variance in each subscale.

Data for the content subscales of the Academic Profile also suggested that there was only one meaningful component underlying the subscales. Eigenvalues of 2.62 and .21 were obtained for data from the Fall quarter, and eigenvalues of 2.62 and .23 were obtained for the data from the Winter quarter. The results presented in the third subtable reinforce the conclusion that a one-component model is sufficient to explain variation in the content subscores on the Academic Profile.

Consistent with results for the other sets of subscales, the relationships among the skill areas of the Academic Profile suggest a unidimensional structure. Eigenvalues of 3.13 and .44 were obtained for the first and second principal components derived from the Fall data, and the eigenvalues for the Winter quarter were 3.14 and .45 respectively. Again, all subscales had significant loadings on the first principal component.

Interrelationships Between the Tests

Given the fact that both the COMP exam and the Academic Profile have a unidimensional structure, the question naturally arises as to whether there is any correspondence between the scores of the two tests. In order to answer this question, students' scores on the two exams were correlated. Only responses from students taking both exams were used (N=38).

Table 4 presents the correlations between the subscales of the COMP exam and the Academic Profile. All of the coefficients in this table are product-moment correlations. Of particular interest is the upper left-hand portion of the table (correlations among content subscores) and the lower right-hand portion of the table (correlations among process/skill subscores).

Table 4:
Correlations Between the COMP Exam and Academic Profile Subscales

ACADEMIC PROFILE	COMP EXAM					
	FSI	UST	UA	COM	SP	CV
HUM	.34	.54	.49	.51	.56	.36
SS	.29	.56	.53	.53	.52	.40
NS	.15	.34	.51	.51	.27	.31
READ	.26	.47	.43	.47	.40	.36
WRIT	.27	.45	.43	.38	.46	.38
CT	.26	.45	.54	.52	.42	.39
MATH	.15	.43	.50	.57	.39	.21

An examination of the upper left-hand portion of Table 4 reveals that all of the content subscores are positively correlated. A closer examination of the correlations among content subscores reveals that the highest correlations among subscales are not for logical counterparts. For example, the Humanities subscale is more highly correlated with Using Science and Technology (.54) than with Using the Arts (.49). Similarly, the Natural Sciences subscale is more highly correlated with Using the Arts (.51) than with Using Science and Technology (.34). Based on these results, it seems safe to conclude that, while the content subscores of the COMP exam and the Academic Profile are interrelated, but there is not a one-to-one correspondence between the subscales.

Establishing a one-to-one correspondence between process/skill subscales is much more difficult because the COMP exam and the Academic Profile differ in what they seem to be measuring. Based on the limited descriptions of the subscales provided by ACT and ETS, no subscale on the Academic Profile corresponds to the Clarifying Values subscale on the COMP exam. In addition, the Communicating subscale on the COMP exam seems to subsume the Academic Profile subscales of Reading, Writing, and Mathematics. It is worth noting that the Mathematics subscale is most highly correlated with Communicating (.57). Two subscales that would be expected to be highly correlated are Critical Thinking and Solving Problems. However, the correlation between these two subscales (.42) is lower than the correlation of Critical Thinking with Communicating (.52) and the correlation of Solving Problems with Writing (.46). Again, there does not seem to be a one-to-one correspondence between process/skill areas on the COMP exam and the Academic Profile.

The absence of a one-to-one correspondence between the subscales of the two exams clearly indicates that these subscales are not interchangeable. However, this does not mean that the two exams, in general, are not measuring the same outcome. It does suggest that the COMP exam and the Academic Profile use slightly different approaches to measuring the same outcome. Indeed, the

presence of a single dimension underlying both exams would help explain the significant correlations between the natural sciences and the arts, and between critical thinking and communicating.

After examining the characteristics of the COMP exam and the Academic Profile, it seems reasonable to conclude that the Academic Profile is slightly superior to the COMP exam in terms of greater item difficulty, greater item variance, and greater scale reliability. The practical result is that the Academic Profile is a better measure of individual differences than the COMP exam. The fact that small score changes can produce very large changes in percentile ranks on the COMP exam also makes the Academic Profile a superior instrument for identifying individual differences.

Despite these indications of the superiority of the Academic Profile, the dominant finding to this point is that neither the subscales of the COMP exam nor the subscales of the Academic Profile are able to differentiate among unique aspects of student outcomes. The presence of a single underlying dimension was confirmed for the content and process/skills subscales of the COMP exam and the Academic Profile. Moreover, there is every indication that the same dimension underlies both exams. Based on these findings, it seems reasonable to recommend against the use of the subscales of these instruments. While a precise identification of the outcome being measured by the two tests must await research on the sensitivity of these instruments to educational experiences, the results obtained thus far suggest that the outcome being measured is very similar to what Spearman (1904) termed "general intelligence."

Sensitivity to Educational Effects

The sensitivity of an assessment instrument to students' educational experiences is an important element in judging the appropriateness of that instrument. Indeed, questions related to educational sensitivity are central to demonstrating the construct validity of an outcomes measure (Cronbach, 1971; Messick, 1987).

In order to evaluate the sensitivity of the COMP exam and the Academic Profile to educational experiences, UTK students' scores on the two tests were correlated with measures of course-taking. Product-moment correlations were calculated for all combinations of coursework variables and test subscales. These correlations are presented in Table 5.

A careful examination of the correlations presented in Table 5 reveals some surprising relationships. Consistent with expectations, humanities coursework was positively related to the Using the Arts subscale on the COMP exam for both the Fall and Winter quarters (.10 and .10). For both quarters, natural sciences coursework (.06 and .09) and mathematics coursework (.10 and .09) were positively related to Using Science and Technology. Similarly, humanities coursework was positively related to the Humanities subscale of the Academic Profile during both the Fall and Winter quarters (.02 and .11 respectively). The combined natural science and mathematics coursework variable was positively related to the Natural Sciences subscale of the Academic Profile for both testing periods (.26 and .19). Contrary to expectations, social science coursework had a negative correlation with Functioning within Social Institutions in the Fall (-.08) and a positive correlation with scores on the subscale

in the Winter (.04). Social science coursework was negatively related to the Social Sciences subscale of the Academic Profile for both the Fall and Winter quarters (-.05 and -.06 respectively).

Table 5:
Correlations of COMP Exam and Academic Profile Scores with Selected Coursework Measures

	COMP EXAM							
	HUMAN	S_SCI	N_SCI	MATHE.	HUMAN	S_SCI	N_SCI	MATHE.
FSI	-.01	-.08	.04	.07	.12	.04	.00	.00
UST	.05	-.18	.06	.10	.09	-.07	.09	.09
UA	.10	-.09	-.03	-.03	.10	-.04	.02	.00
COM	-.03	-.16	.03	.16	.02	-.07	.05	.13
SP	.14	-.04	-.01	-.05	.02	-.07	.05	.13
CV	.05	-.14	.04	.02	.11	-.01	.01	.01
TOTAL	.06	-.15	.03	.06	.10	-.03	.05	.04

	ACADEMIC PROFILE					
	HUMAN	S_SCI	M/N_SCI	HUMAN	S_SCI	M/N_SCI
HUM	.02	-.02	.07	.11	-.04	-.09
SS	-.08	-.05	.12	.03	-.06	-.03
NS	-.14	-.19	.26	.02	-.23	.19
READ	.03	-.01	.25	.08	-.05	-.07
WRIT	-.04	-.04	.08	.07	-.07	-.11
CT	-.04	-.09	.14	.07	-.04	.02
MATH	-.21	-.20	.29	-.02	-.28	.27
TOTAL	-.07	-.10	.16	.06	-.12	.03

For the process/skill subscales, natural science and mathematics coursework were positively related to Communicating, both in the Fall (.03 and .16) and in the Winter (.05 and .13). The combined mathematics and natural science coursework variable was positively related to the Academic Profile's Mathematics subscale for both quarters (.29 and .27). The combined mathematics and natural science coursework variable also was positively related to Critical Thinking during the Fall and Winter quarters (.14 and .02 respectively). Interestingly, natural science coursework and mathematics coursework tended to be negatively related to Solving Problems on the COMP exam. For both the Fall

and Winter quarters, humanities coursework was positively related to Solving Problems (.14 and .12 respectively).

The relationships identified in Table 5 provide a clear indication that neither the COMP exam nor the Academic Profile is particularly sensitive to students' educational experiences (at least as they are measured by patterns of coursework). Indeed, the key to enhanced performance on both the COMP exam and the Academic Profile is not to be found in a broad general education. Instead, the key to improve test performance is a specialized pattern of course taking that deemphasises social science coursework and emphasizes mathematics, natural science, and to a lesser extent, humanities coursework.

One explanation for the relationships among coursework measures and test scores is that some outside variable is producing spurious relationships. Previous research has found that students' entering levels of academic ability influence coursework and test performance. The net effect is that entering ability distorts the relationship between coursework and test performance (Pike, in press).

In order to determine if academic ability influences performance on the COMP exam and the Academic Profile, students' scores were correlated with their entering ACT Assessment scores. These results are presented in Table 6.

An examination of the correlations presented in Table 6 reinforces the view that entering academic ability is a major determinant of student performance on outcomes measures like the COMP exam and the Academic Profile. As the data in Table 6 clearly show, all of the subscales on the COMP exam and the Academic Profile are significantly related to students' ACT Assessment scores. For the Fall quarter, the range of correlations is from .35 (Using the Arts) to .51 (Communicating) on the subscales of the COMP exam, and from .59 (Reading) to .70 (Mathematics) on the subscales of the Academic Profile. For the Winter quarter, the range is from .39 (Solving Problems) to .50 (Communicating) on the COMP exam, and from .56 (Writing) to .67 (Social Sciences) on the Academic Profile. ACT Assessment scores are significantly correlated with total scores on the COMP exam (.57 and .56) and on the Academic Profile (.73 and .69) for both the Fall and Winter quarters.

While the correlations between ACT Assessment scores and COMP scores are somewhat lower than comparable correlations for the Academic Profile, this is not a cause for celebration. At least part of this difference is due to lower levels of reliability for the COMP exam.

Table 6:
Correlations of COMP Exam and Academic Profile Scores
with Entering ACT Assessment Scores

	COMP EXAM	
	Fall	Winter
FSI	.49	.43
UST	.50	.47
UA	.35	.41
COM	.51	.50
SP	.40	.39
CV	.42	.41
TOTAL	.57	.56

	ACADEMIC PROFILE	
	Fall	Winter
HUM	.67	.61
SS	.69	.67
NS	.69	.66
READ	.59	.60
WRIT	.60	.56
CT	.69	.65
MATH	.70	.65
TOTAL	.73	.69

Conclusion

Taken as a whole, the present research found very little that would recommend either the COMP exam or the Academic Profile as outcomes measures. Indeed, the one clear finding to emerge from this research is that both the COMP exam and the Academic Profile are unidimensional measures that are not particularly sensitive to students' general education experiences. As a result, institutions would be unwise to use the subscales of the COMP exam or the Academic Profile to evaluate the components of a curriculum. Use of these scales to suggest curriculum changes could even be counterproductive.

In sum, the COMP exam and the Academic Profile represent measures of individual differences. More specifically, these tests are very powerful measures of general academic ability. Because of their sensitivity to individual differences, the COMP exam and the Academic Profile are not appropriate for evaluating the impact of general education programs.

References

American College Testing Program. (1987). College Outcome Measures Program. Iowa City: Author.

American College Testing Program. (1988). [Senior sample means on the COMP exam]. Unpublished data table.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 443-507). Washington, D.C.: American Council on Education.

El-Khawas, E. (1987). Colleges reclaim the assessment initiative. Educational Record, 68, 54-58.

ETS College and University Programs. (1988). The Academic Profile: Information booklet. Princeton, NJ: Educational Testing Service.

Ewell, P. T. (1987) Assessment, where are we? Change, 19, 23-28.

Forrest, A., & Steele, J. M. (1982). College Outcome Measures Project: Defining and measuring general education knowledge and skills. Iowa City: American College Testing Program.

Gulliksen, H. (1950). Theory of mental tests. New York: John Wiley & Sons.

Messick, S. (1987). Validity (Report no. RR-87-40). Princeton, NJ: Educational Testing Service.

Pike, G. R. (in press). Students' background characteristics, educational experiences, and educational outcomes: A model for evaluating assessment instruments. Proceedings of the conference on developing assessment partnerships between community and senior colleges: The theory and practice of outcomes assessment. Fairfax, VA: George Mason University.

Spearman, C. (1904). "General intelligence," objectively determined and measured. American Journal of Psychology, 15, 259-292.