DOCUMENT RESUME

ED 236 209 TM 830 720

AUTHOR Peterson, Gary W.
TITLE The Development of a Baccalaureate Outcome Measure Based on a Generic Skills Theory of Human Performance.
PUB DATE 11 Apr 83
NOTE 48p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Montreal, Quebec, April 11, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Bachelors Degrees; Cognitive Processes; Difficulty Level; *Essay Tests; Higher Education; Interrater Reliability; *Outcomes of Education; Problem Solving; Program Evaluation; *Rating Scales; *Test Construction; Test Reliability

ABSTRACT

Even though several national testing firms have developed measures to evaluate the effectiveness of baccalaureate education, there continues to be a general reluctance on the part of faculty in colleges and universities to accept these measures as criteria on which to evaluate educational programs. Some of the resistance appears to lie in the lack of validity of commercially prepared tests as measures of their competencies (often 50 to 100 in number), the amount of cost and effort to evaluate these competencies with the kinds of performance tests they seem to prefer, and questionable psychometric rigor of locally made performance tests. A series of three prototypical real-life problem solving exercises and rating criteria were developed to address these limitations. A theory of human performance provides a rationale for the development of the generic skills test. The stimuli are in the form of real-life scenarios, and require short-answer essay responses. Rating scales use descriptions of high, medium and low performance attributes to evaluate responses. The results of a field test indicate that the difficulty level of exercises appears acceptable, the problem-solving constructs are independent and meaningful, and that the interrater and alpha reliabilities are sufficiently high to warrant its use as a program evaluation instrument. The contribution of basic intellectual factors, maturation and educational experiences to generic performance should be ascertained through further research. (Author)

The Development of a Baccalaureate Outcome Measure Based on a Generic Skills Theory of Human Performance[1]

by

Gary W. Peterson
Florida State University

Abstract

Even though several national testing firms have developed measures to evaluate the effectiveness of baccalaureate education, there continues to be a general reluctance on the part of faculty in colleges and universities to accept these measures as criteria on which to evaluate educational programs. Some of the resistance appears to lie in the lack of validity of commercially prepared tests as measures of their competencies (often 50 to 100 in number), the amount of cost and effort to evaluate these competencies with the kinds of performance tests they seem to prefer, and questionable psychometric rigor of locally made performance tests. A series of three prototypical real-life problem solving exercises and rating criteria were developed to address these limitations. A theory of human performance provides a rationale for the development of the generic skills test. The stimuli are in the form of real-life scenarios, and require short-answer essay responses. Rating scales use descriptions of high, medium and low performance attributes to evaluate responses. The results of a field test indicate that the difficulty level of exercises appears acceptable, the problem-solving constructs are independent and meaningful, and that the interrater and alpha reliabilities are sufficiently high to warrant its use as a program evaluation instrument. The contribution of basic intellectual factors, maturation and educational experiences to generic performance should be ascertained through further research.

[1]Paper presented at National Council on Measurement in Education Annual Convention, Montreal, Quebec, April 11, 1983.

The Development of a Baccalaureate Outcome Measure Based on a Generic Skills Theory of Human Performance

by

Gary W. Peterson
Florida State University

While several showcase competency-based liberal arts programs have been implemented in higher education and several national testing firms are either developing or marketing instruments to measure baccalaureate outcomes (Forrest and Steele (1981), Warren, (1980), Winter, McClelland and Stewart 1981), the nature of the baccalaureate outcomes and their measurement is still very much open to question. Further, the press for accountability in higher education (Peterson and Stakenas, 1981), has resulted in many institutions either trying purchase outcome measures from commercial vendors or trying to develop their own in order to demonstrate "value-added" contributions to intellectual skills (Katz, 1982, Astin 1982). In spite of the existence of model competency programs and commercially prepared instruments faculty often reject available tests out of hand at face value or fall short in their efforts to create them (Peterson and Watkins, 1979, Peterson, 1982).

There are additional deterants that discourage faculty from using competency outcome measures to evaluate the effectiveness of baccalaureate programs. First, faculty, in the process of defining desired baccalaureate outcomes, often formulate ponderous lists of 50 to 100 competency statements. But, without a clear and coherent theoretical framework with which to abstract such lists, the selecting or developing of

assessments is confusing, if not overwhelming. Secondly, while faculty are generally favorably disposed to the use of free response questions as a strategy for assessing higher order thinking skills (Peterson and Watkins, 1979), the amount of time, inconvenience and resources required to administer and score such instruments also present drawbacks--not to mention certain methodological problems. Thirdly, the task of developing a baccalaureate test that takes into account differences among cohorts of students in terms of intellectual factors as well as educational and life's experiences also presents formidable psychometric challenges. A prototype baccalaureate outcome measure was developed from a comprehensible theoretical framework that offers practical utility and psychometric integrity.

Based on earlier theoretical research (Peterson 1981, 1982, Peterson and Rumsey, 1981), the prototype baccalaureate measure was derivied from a model of human performance that links baccalaureate skills dynamically (see Travers, 1980) so that multiple skill dimensions can be assessed through the use of a single stimulus situation (i.e., scenario). To assist faculty in developing locally-made versions of the instrument, a framework for the development of scenarios and their respective stimulus and response attributes is also presented. Analyses pertaining to construct validity, reliability and utility are included.

A generic skills theory of human performance

The development of the prototype baccalaureate assessment battery began with the concept of competence. Competence may be considered as the integration of knowledge, skills, and attitudes required to perform adult

roles effectively (McClelland, 1973; and Goleman, 1981). A competent individual, then this one who is able to meet effectively and efficiently the demands of occupational, personal, and community roles as in the case of a competent physician, army officer, or housewife. Competent performance is based on the mastery and integration of component cognitive and psychomotor skills and knowledge. These component skills are typically labeled competencies.

Some skills, however, seem to occur over and over again in the analysis of many and diverse occupational or life tasks. These pervasive and transcendant skills are designated as generic skills. According to Woditsh (1977), the term, generic, connotes a function or a pattern of activity that is recurrent in a wide series of discrete purposive behaviors. "Generic skills are basic in the sense that they are ubiquitious: they show up again and again as components or instances of successful behavior" (pg. 8). Thus one infers the presence of generic skills from the successful performance of discrete tasks which constitute diverse occupational or life roles. A task is defined as a unit of behavior with a beginning and ending point that can be accomplished in hours rather than days or weeks (Branson, 1980).

Important attributes of generic skills may include the following (Peterson, 1981):

- Generic skills undergrid what is commonly referred to as generalized problem-solving behavior;
- A generic skill is an ability or capability that possesses its own unique hierarchy of discrete related component skills;
- A generic skill is pervasive and recurs across academic or professidonal disciplines of study and even life or job skills;

- The mastery of knowledge base underlies the development and demonstration of generic skills;
- The demonstration of generic skills requires the integration of discrete lower order component skills and knowledge; and
- Individuals who have mastered generic skills are able to apply them in a variety of real-life situations or contexts to solve problems encountered in adult roles in society.

Thus, generic skills are constructs which could be viewed as "common denominators" of learning among various disciplines of study and may be used to compare student development across academic programs within a university or even among colleges and universities. Stated in more formal terms, generic skills are constructs of intellectual ability that capture common variance across-discretely different discipline-oriented performance tasks.

Structural and functional relationships among generic skills

Four generic skills have been identified from previous research by Peterson (1982) and Warren (1980) for the development of measures: they include Communication, Disciplined Inquiry, Valuing, and Execution. These skills undergird a capstone skill referred to as generalized Problem-solving. Structurally, according to the author's formulations (see figure 1, each of these skills in turn possesses its own unique hierarchy of subordinate skills such as the perceiving and interpreting of verbal, non-verbal and quantitative symbols in Communications or analyzing or synthesizing in Disciplined Inquiry. Further, the capacity to demonstrate generic skills is based on a set of basic cognitive abilities (e.g., relationships among words and concepts, proportionate logic, spacial

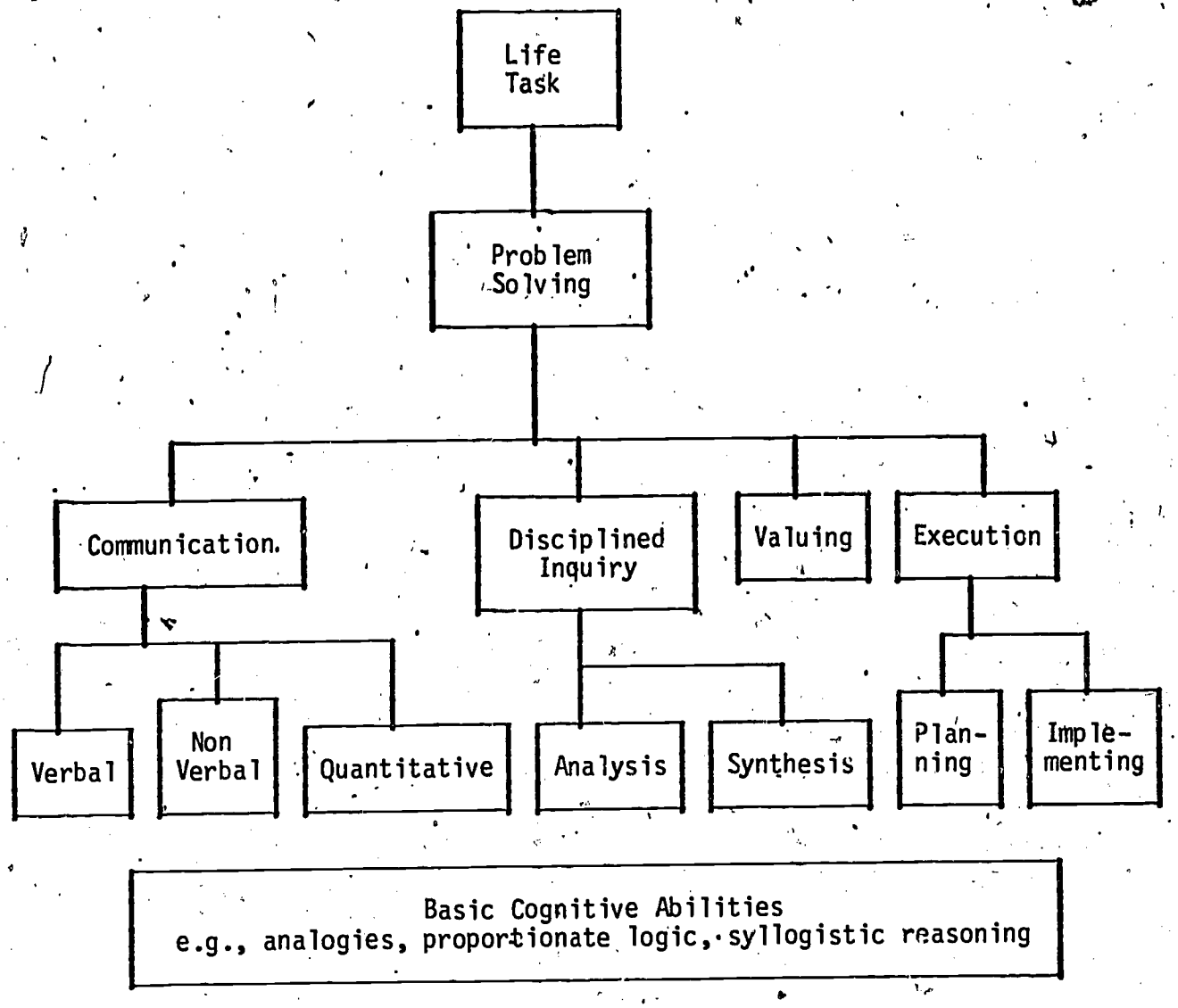
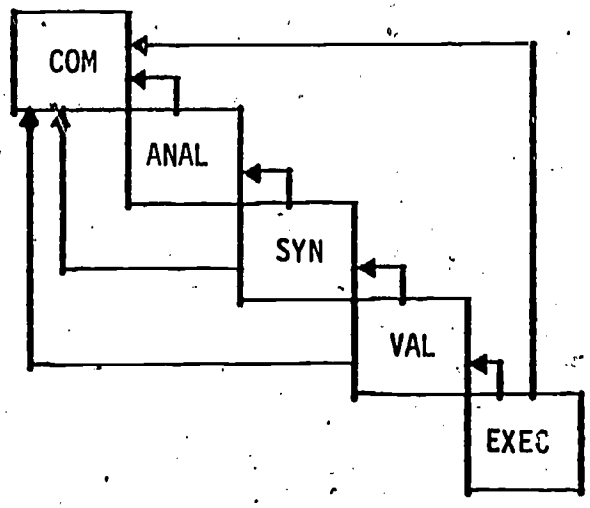


Figure 1. An hypothetical structure and relationships among problem solving, generic, and basic cognitive skills.

relationships) that are typically measured by common intelligence and aptitude tests.

In terms of the functional characteristics of the model, the first step in coping with problem situations (anomalies, needs, dissonances) is to perceive and interact with the phenomenon of inquiry (i.e. a problem solving task). This calls for receiving, interpreting, and sending verbal, nonverbal, and quantitative messages (Communication). Individuals must then be able to inquire into the phenomenon by reducing it into its component parts, and relating them to each other (Analysis). They must then extract and organize information to formulate potential courses of action to reduce or eliminate the anomaly (Synthesis). Next, each alternative course of action must be evaluated in terms of costs and benefits to themselves, their social or occupational groups, and to the wider society in terms of higher order principles or ethics (Valuing). Finally, individuals must structure activities to achieve a goal by defining objectives in a coherent means-ends relationships (Planning) and carry out the plan (Implementation).

All of these capabilities are brought to bear in performing essential life and occupational tasks in which the cue functions are ambiguous, the alternative courses of action not predetermined, and the consequences of implementing various courses of action uncertain. If there is a lack of development in any of these capabilities, a person's potential as a problem-solver and hence performer of complex life tasks is limited concomitantly. The portrayal of this process may be depicted in Figure 2 on the next page in the form of an N^2 chart. One can note from the chart how successive skills convolute on preceding skills. The engaging of a latter



Where:

COM is perceiving a problem or need (Communication)
ANAL is relating components (Analysis)
SYN is formulating possible courses of action (Synthesis)
VAL is estimating costs and benefits for each course of action (Valuing)
EXEC is planning and implementing a solution (Execution)

Figure 2. N^2 Chart for Interrelationships Among Subordinate Generic Problem Solving Skills

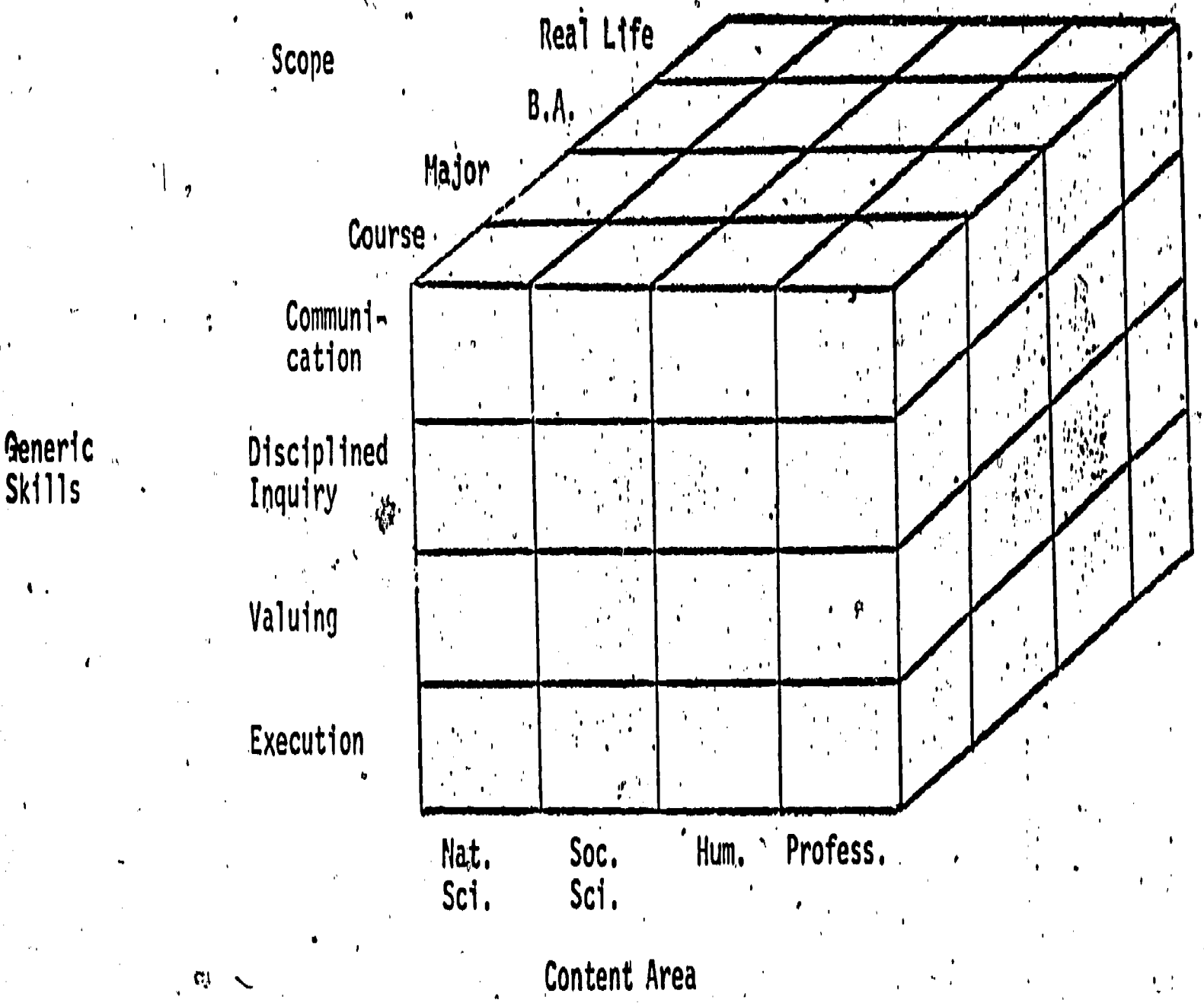


Figure 3. A model for the relationship among skill, content, and scope in generic skill test development.

a narrow or discrete sphere of knowledge, a mastery of several operations and an ability to apply these skills in familiar instructional problem situations. Thus the cue function (i.e. test item stimulus) should be recognizable to students who have had the course and the responses may require only the retrieval of knowledge and skills practiced in familiar course contexts.

A second or more distal level of outcome, i.e. the major or minor field, concerns whether students are able to demonstrate generic skills using a wider range of knowledge and more complex operations and procedures than at the proximal level. Students must be able to integrate knowledge and skills from a series of related courses and use specialized knowledge to address more complex situations than at the course level of outcome. Assessment situations at the second level also involve more complex cue functions than at the proximal level. Students must be able to perceive order in more ambiguous stimulus situations by being able to to differentiate relevant from irrelevant information. They must bring to bear more complex procedures than at the course level, select appropriate operations according to their utility and efficiency, and integrate them to address higher order purposes. The responses may also be more sophisticated requiring more complex and ordered sequences of expression while using appropriate information to document and to justify. These capabilities are performed through the "gestalt" of a discipline of study.

Assessment of generic skills at the third level, i.e., the B.A. level, requires students to investigate issues or problems integrating both specific major field and broad general education perspectives. A student should be able to not only use formal structures from one's discipline of study,

but also to be able to draw upon other disciplines encountered through general education and minor field concentrations. For example, if biology majors are investigating whether a bridge should be constructed across an estuary, they should be able to determine the potential impact such a structure could have on the flora and fauna of the local ecosystem as well as to draw from the Humanities and Social Sciences to address possible economic, social, and aesthetic implications.

The cue function of an assessment task at the third level also involves imposing order and more structure on more ambiguous situations that at the proximal or second level. A wider repertoire of cognitive skills is also required to perform tasks at this level and responses entail the integration of a wider diversity of principles and supportive detail. Students must also be aware of the biases and values of the respective disciplines as well as be congnizant of their own personal values. They should be able to cope with value conflicts by applying a higher order of recognized sets of principles and ethics.

The fourth or ultimate level generic skill assessment concerns evaluating whether students are able to "break" the boundaries of disciplinary thought as acquired in sequences of academic courses and are able to demonstrate the ability to use generic skills to cope with real life issues and problems such as choosing a political candidate, selecting a career, resolving a personal conflict, making decisions in professional situations, or taking a stand on a political issue. In such assessments, students are presented with a novel situation and are required to structure its apparent elements, seek appropriate information, formulate rules when necessary, and to distinguish among personal, cultural, and disciplinary

values. At this level, the assessment tasks may well appear to be "discipline-free" in the sense that responses do not require the recalling facts, definitions, concepts or principles from a single academic discipline. Students are required to perform a more difficult task--to seek out, select and evaluate information pertaining to an issue, a problem, or pending decision.

At this ultimate level of "scope" dimension, the stimuli of the assessment tasks tend to be ambiguous and unstructured. Students are asked to assume a socially responsible role presented with an issue, a dilemma, a question, a problem to solve or a decision to make. The task should be sufficiently complex so as to require the integration of a variety of disciplinary perspectives and skills. Students from the respective disciplines are likely to demonstrate differences in the manner in which they respond to problem-solving situations because the perspectives and analytical "tools" (Communications, Disciplined Inquiry, Valuing and Execution) they have acquired will reflect the kinds of educational experiences they have had previously. All three exercises in the prototype test assess generic skills at the ultimate level, each drawing upon perspectives from a variety of disciplines.

Assumptions

Three assumptions or propostions are made in the development of the prototype baccalaureate outcome measure:

- There is a continuum of cognitive complexity that underlies each generic skill dimension which is reflected in terms of high, medium and low levels of performance;
- An aim of higher education is to assist students in developing problem solving capabilities which necessarily includes important

component skills (i.e. generic skills);

- One criterion with which to evaluate academic program effectiveness should include an assessment of growth in generic skill capabilities of students.

Method

Test Development

The scenarios. Three scenarios were developed using the Skill X Content X Scope model depicted in Figure 3. The first exercise, Cuban Crisis, was designed to emphasize Humanities and Social Science perspectives of the cube; the second, Land Development, was designed to draw on the Natural and Social Sciences; and third, Token Economy draws on Social Sciences and Professional education. Cuban Crisis was developed with the assistance of a faculty colleague in the Department of Religion; Land Development was developed by the author from a series of articles and editorials in a local newspaper; and Token Economy was also formulated by the author. Each of the scenarios requires an individual to assume a social role such as a social caseworker (Cuban Crisis), a county commissioner (Land Development), and a county school board member (Token Economy). The scenarios were intended to create interesting and engaging situations to help sustain motivation in the task. The situations in the scenarios were designed to be sufficiently complex so as to compel high levels of performance for of each of the generic skills. The three scenarios contained between 150 and 250 words and are located in Appendix III.

The directive statements. The directive statements are the keys to the effectiveness of each task in terms of eliciting maximum performance in

the respective generic skills. The directives were derived from liberal arts competency statements obtained from earlier research efforts (Peterson and Watkins 1978 and Peterson, 1982). They are the same for each scenario and are as follows:

1. Problem Solving

   Subskill: Formulate a plan of action to rectify a need or anomaly based on a higher order, principled rationale.

   Directive Statement: Describe which course(s) of action you would recommend in the above situation and explain why. Please elaborate as much as you can and use the back of the paper if necessary.

2. Communication

   Subskill: Comprehend complex situations by expressing the main idea in one's own words.

   Directive: As a (role specified) what is the central issue for you in the above scenario?

3. Analysis

   Subskill: Perceive elements in a problem situation by demonstrating the ability to view an event from multiple perspectives.

   Directive: Describe the problem from A's perspective, B's perspective or C's perspective.

4. Synthesis

   Subskill: Formulate a variety of discretely plausible courses of action given a problem situation.

   Directive: List as many possible courses of action that might be taken in the above scenario, including options you would not choose to follow. Use the back of the page if there is not enough space for all of your alternatives.

5. Valuing

   Subskill: Infer values from behavior

Directive: What values underlie A's behavior, B's behavior, and C's behavior?

6. Execution

Subskill: Perceive logical means-ends relationships with sufficient detail.

Directive: For the solution you recommended in Part I (Problem Solving), outline the sequence of actions that you would take to implement the solution. Describe what you would do first, second, third, etc.

The rating scales. The development of rating scales involved the delineation of performance attributes for high (5), medium (3) and low (1) anchor points along a 5-point continuum. The works of developmental psychologists such as Piaget (See Flavel, 1963), Kohlberg (1971) and Perry (1970) were helpful in developing a priori general conceptualizations of the respective performance continua. The description of performances were developed and refined on the basis of a pilot test. The scenario and the directives for Cuban Crises (Exercise 1) are presented in Appendix I. The rating scales appear in Appendix II.

Test battery assembly. The prototype baccalaureate test consists of a background information form, a vocabulary test, and three generic skills exercises. A 5-point rating scale was appended to Cuban Crisis and Land Development to assess the degree of familiarity with the content of the respective scenarios. The background information form asked for the subjects' age, total number of semester hours completed, the number of semester hours completed in their major field and the number of semester hours completed in Psychology. The vocabulary test was the Wide Range Vocabulary Test (1962), a 12-minute test to gain an indication of a general intelligence factor (Spearman g) that ostensibly underlies the demonstra-

tion of generic skills according to the model, This short vocabulary test can be used in program evaluation purposes as a covariate to equate performances among cohort groups in terms of verbal ability. Each exercise consists of two parts: (1) the decision (Problem Solving) and (2) the inquiry (the assessment of generic skills). The first part, the decision, seeks to establish a goal and rationale for coping with a problem situation while the second part, the inquiry, seeks to assess the level of generic skill proficiency undergirding the decision. The procedure is somewhat analogous to the procedures used to administer psychological projective instruments.

Pilot testing. The pilot test subjects consisted of 17 juniors and senior from an upper division Human Relations and Communications course offered in the College of Education and 5 graduate students in Counseling. Both groups completed the three exercises as homework assignments. As a result of the pilot test, both the directive statements and model responses at the 5, 3 and 1 anchor points were revised.

Field-Test Administration

Samples. Three groups of students were sought for the field test: lower division, upper division and graduate students to gain a full range of generic skill performances and background characteristics with-which to analyze relationships among constructs. The lower-division students were volunteers from Introductory Psychology classes who received course credit for participation in faculty research. The upper division students were paid volunteers ($10 per session) who came primarily from ROTC classes and

Human Relations and Communication classes. Students from almost all majors attend these classes. The graduate students were volunteers from the authors's graduate classes in the College of Education, most of whom were in their first year of graduate school. The sample consisted of 20 freshmen and sophomores, 26 juniors and seniors, and 16 graduate students. There were 27 males and 35 females. Major field representation was 61% Professional Schools, 25% Social Sciences, 5% Natural Sciences and 5% Undecided. The author recognizes that a stratified random sampling procedure would have been preferred to offset the risk of sampling error bias.

Testing sessions. When visiting classes to secure volunteers, a sign-up sheet was circulated for subjects to attend scheduled testing sessions. Subjects were informed that a testing session would take between 2 and 3 hours. The show-rate for Psychology students and graduate students (non paid valunteers) was 75% while only 50% of the upper division students (paid volunteers) appeared for testing after signing up.

When subjects arrived for the testing sessions, they first completed the background information questionnaire and then the 12-minute Wide Range Vocabulary Test. They were then informed they had unlimited time to complete the three generic skills exercises. A brief description of a token economy was given orally by the proctor since some of the subjects were unfamiliar with what a token economy was. Most students completed the exercises within two hours and no student took longer than three. The order in which the exercises were administered was Cuban Crisis, Land Development and Token Economy.

Rating the Responses

Each response was rated three times by three ABD doctoral students in Counseling. Before beginning the rating procedures, they were informed about the nature of the project, the skills and the scales. They were instructed, for each of the dimensions, to first rate the responses of the first ten subjects and then to return to the beginning and readjust any scores following a second reading. Then, following a question and answer period, the remaining 52 responses were scored. Approximately seven hours were required of each rater to score 62 samples for all six generic skill dimensions for each exercise. Thus about 21 hours of rating time was required for each rater to evaluate all three exercises for 62 examinees. The raters evaluated all protocols one dimension at a time to eliminate halo effects across skills. The tests were sequenced in alphabetical order and the identity of subjects remained anonymous to the raters. To offset potential ordering effects, one rater scored a given set responses forward, another backward and a third began in the middle and proceeded forward.

Results

Test difficulty. The means and standard deviations and range for each exercise as well as the scores for all three exercises combined are presented in Table 1. All of the means were near the midrange of the scales (i.e. medium level responses) except for the Valuing dimension which was between the medium and low range (2 level responses). All of the distributions met normality criteria for skew and kurtosis. The complete range of performances (lowest possible (3) and highest possible (15)) were found in twelve of eighteen tasks (three exercises X six skills). The ranges of

Table 1

Difficulty level, range and reliability

Skills / Exercises	Mean	S.D.	Maximum Possible Scale	Observed Range	Alpha Reliability	Mean Interrater Reliability
Problem Solving						
Exercise 1	8.18	2.85	15	3-15	.79	.56
Exercise 2	8.15	2.78	15	3-15	.78	.54
Exercise 3	8.08	2.87	15	3-15	.78	.55
Combined	24.40	6.44	45	13-41	.81	
Communication						
Exercise 1	7.52	2.93	15	3-13	.79	.57
Exercise 2	9.45	3.52	15	3-15	.86	.67
Exercise 3	7.90	2.94	15	3-15	.75	.50
Combined	24.87	7.41	45	10-43	.84	
Analysis						
Exercise 1	9.00	2.57	15	4-14	.63	.35
Exercise 2	8.74	1.60	15	3-13	.67	.46
Exercise 3	8.10	2.82	15	3-15	.73	.48
Combined	24.84	5.20	45	15-40	.73	
Synthesis						
Exercise 1	7.34	2.33	15	3-13	.67	.41
Exercise 2	9.65	2.82	15	3-15	.82	.60
Exercise 3	8.73	3.09	15	3-15	.80	.56
Combined	25.71	6.17	45	15-41	.80	
Valuing						
Exercise 1	5.81	2.81	15	3-15	.82	.60
Exercise 2	6.56	2.34	15	3-13	.69	.44
Exercise 3	6.32	3.03	15	3-15	.88	.71
Combined	18.69	6.47	45	9-33	.85	
Execution						
Exercise 1	7.00	2.41	15	3-13	.66	.41
Exercise 2	8.42	2.64	15	3-15	.68	.43
Exercise 3	8.74	3.27	15	3-15	.89	.73
Combined	24.16	6.12	45	12-39	.78	

Exercise 1 = Cuban Crisis
Exercise 2 = Land Development
Exercise 3 = Token Economy

scores of the combined scales created by summing nine ratings across three exercises were from 24 to 33. In only one skill, Valuing, did a student receive nine, 1-ratings across all three exercises. No student earned all 5's across all the exercises for any of the skills. One student in Analysis earned seven, 5 ratings and two, 4 ratings. Therefore, the levels of difficulty, ranges of scores and variability of scores within generic skill dimensions appear to be sufficient for use as a program evaluation measure.

Test reliability. The reliabilities (coefficient alpha) for the scale of 18 tasks (3 exercises x 6 skills) are also presented in Table 1. The alphas range from .63 to .89 with a mean of .71. The alphas for the combined scales formed by adding nine ratings for each subject ranged from .73 to .85 with a mean of .80. The interrater product-moment correlations for each of the 18 individual tasks ranged from .35 to .73 with a mean of .53. This level of interrater agreement is typical of measures using high levels of inference to rate the responses. The alpha reliabilities of the combined ratings were viewed as sufficient for use of the instrument in program evaluation.

The relationships between exercises. Both zero-order and attenuated correlations of skills across exercises are shown in Table 2 on the next page. The range for zero-order correlations is .19 to .49 with a mean of .37. Likewise the range for the attenuated correlations is .26 to .65 with a mean of .49. These coefficients represent an index of the degree of commonality of skills across exercises. Using the attenuated coefficients to estimate the relationship between two variables assuming perfect measures, the average amount of common variance between two identical skills across

Table 2

Zero-Order and Attenuated Correlations by Skills Between Exercises

Skills / Exercises	r_{xy} (r_{tt}) Ex2	r_{xy} (r_{tt}) Ex3	Skills / Exercises	r_{xy} (r_{tt}) Ex2	r_{xy} (r_{tt}) Ex3
Problem Solving			Synthesis		
Exercise (1)	.31 (.40)[a]	.44 (.57)	Exercise (1)	.19 (.26)	.36 (.49)
Exercise (2)		.32 (.41)	Exercise (2)		.43 (.58)
Communication			Valuing		
Exercise (1)	.49 (.60)	.49 (.65)	Exercies (1)	.49 (.65)	.36 (.42)
Exercise (2)		.33 (.41)	Exercise (2)		.48 (.62)
Analysis			Execution		
Exercise (1)	.37 (.57)	.33 (.49)	Exercise (1)	.37 (.55)	.37 (.48)
Exercise (2)		.24 (.34)	Exercise (2)		.20 (.26)

a. Numbers in parentheses are correlations for attenuation based on alpha reliabilities.

different exercises is 24%. With zero-order correlations, the average amount of common variance between skills decreases to 14%. Further, even though the exercies were designed to represent different content domains, on the skill x content x scope matrix, the range of strength of attenuated intercorrelations within skill areas appears small (maximum .22 between highest and lowest for Synthesis). Therefore, because of the modest correlations (i.e. shared variances) across exercises within skills, entities called generic skills were measured.

The relationship among generic skills. The matrix for both zero-order and attenuated correlations among generic skills is presented in Table 3. The generic skill scores were determined by summing the nine ratings across three exercises (three ratings for each of three exercises). The range of zero-order correlations among skills is .33 to .66 with a mean of .55 while the range for attenuated correlations is .41 to .84 with a mean correlation of .69. Therefore, using the mean attenuated correlation among all generic skills, there is an average of 48% common variance between individual skills. While this may seem high at first glance, one must be mindful that there is also an average of 52% unique variance between the skills. Thus, the six individual skill dimensions may be considered as independent and meaningful constructs.

Familiarity with problem situation. An investigation was made regarding the degree to which familiarity with the problem scenario in real-life situations (Cuban Crisis and Land Development) or the extent to which of coursework in a technical situation (Token Economy) might be related to generic skill performance. In the two Real-Life exercises,

Table 3

Zero Order and Attenuated Correlations Among Generic Skills

Generic Skills	Com r_{xy} (r_{tt})	Anal r_{xy} (r_{tt})	Syn r_{xy} (r_{tt})	Val r_{xy} (r_{tt})	Exec r_{xy} (r_{tt})
Problem Solving	.61 (.74)[a]	.63 (.82)	.63 (.78)	.60 (.72)	.49 (.62)
Communication		.52 (.66)	.50 (.61)	.45 (.53)	.33 (.41)
Analyses			.60 (.79)	.66 (.84)	.63 (.84)
Synthesis				.63 (.77)	.48 (.61)
Valuing					.49 (.60)

a. Numbers in parentheses are correlations corrected for attenuation based on alpha reliabilities.

Cuban Crisis and Land Development, subjects rated their level of familiarity with the two situations on a five point scale (see Appendix II). In the third exercise, Token Economy, subjects were asked to record the total number of semester hours completed in Psychology. The correlations between responses to the familiarity scales and the first two exercises as well as credit hours taken in Psychology with respect to the third exercise are presented in Table 4. The results indicate there was only one exercise within one skill in which there was a significant correlation between the familiarity scale and generic skill performance. This one significant correlation was considered to be an artifact of the data. Therefore it was concluded that familiarity with an event in real-life scenarios in this battery has little if no bearing on the demonstration of generic skills.

Significant correlations were however obtained between credit hours in psychology and generic skill scores in the Token Economy Exercise. When interpreting these results, it is well to refer to Table 5 on page 26 in which correlations between generic skills and a variety of background factors are presented. The strength of correlations between generic skills and age and total credit hours completed appears strong. Thus the relationship between credit hours taken in Psychology and generic skill scores may be more owed to maturation, total educational experience and verbal ability than to mastery of content in psychology. On the basis of these data the conclusion cannot be made in terms of whether the mastery of specific content is related to generic skill performance. Thus the relationship between the degree of content mastery and generic skills is still an open issue. The implications of this finding for academic program evaluation are discussed later.

Table 4

Correlations of Familiarity With Problem Situation or Content Domain

Skills Exercise	r	Skills Exercise	r
Problem Solving		Synthesis	
Exercise 1	.03[a]	Exercise 1	-.01
Exercise 2	-.06[b]	Exercise 2	-.08
Exercise 3	.49***[c]	Exercise 3	.31**
Communication		Valuing	
Exercise 1	-.06	Exercise	-.06
Exercise 2	.06	Exercise	-.10
Exercise 3	.28*	Exercis	.38***
Analysis		Execution	
Exercise 1	.22*	Exercise 1	.00
Exercise 2	-.09	Exercise 2	-.03
Exercise 3	.52***	Exercise 3	.45***

a) Correlation with 5 point familiarity scale with Cuban Culture
b) Correlation with 5 point familiarity scale with Land Development issues
c) Correlation with total semester hours completed in Psychology

* $p < .05$
** $p < .01$
*** $p < .001$

Relationship between generic skills and verbal ability, age and educational experience. Correlations between age, elements of educational experience and vocabulary are presented in Table 5. The mean correlations between generic skill performance and age is .56, year in college .52, total semester hours .52, semester hours in major .49, semester hours in Psychology .49, semester hours of general education .34 and vocabulary .50. Apparently, educational experience, maturation and intellectual factors are all related to generic skill performance. Unfortunately all of these factors are themselves highly interrelated in so that the determination of their relative contribution to generic skill performance could not be ascertained with regression analyses due to the presence of multicolinearity (Lewis-Beck, 1980). If the subject pool had included a group of adult learners with limited educational experience, then perhaps some light could have been shed on the nature-nurture controversy with respect to the demonstration of generic skills. Nevertheless these results point out that the use of a general intelligence measure should be included as a covariate when comparing generic skill performance between or among student cohorts in academic program evaluation.

Discussion

A prototype baccalaureate outcome measure was developed to make the assessment of higher-order thinking skills more theoretically coherant, practical, and psychometrically defensible. A problem solving model was advanced in which five generic skills were identified which undergird the capstone skill, Problem Solving, namely Communication, Analysis, Synthesis, Valuing and Execution. An assessment strategy was advanced in which all of

Table 5

Generic skills Correlations with Background Variables.

	Problem Solving	Com.	Anal.	Syn.	Val.	Exec.
Age	.65	.37	.60	.61	.56	.51
Year in College	.54	.31	.58	.58	.56	.53
Total Semester Hours Completed	.56	.35	.57	.60	.57	.49
Semester Hours in Major	.54	.33	.55	.56	.53	.46
Semester Hours in Psychology	.53	.27	.60	.43	.49	.60
General Education	.34	.22	.34	.39	.38	.34
Vocabulary	.45	.31	.63	.55	.62	.46

these skills may be treated in the context of one problem solving situation. The following discussion focuses on the validity, reliability and utility use of the measure in academic program evaluation.

Validity as a generic skill measure. At the present stage in the development of generic skills measures, a major concern regarding validity is from the standpoint of construct validity. Data were analyzed with respect to the independence of dimensions, commonality of skills across tasks, the contribution of intelligence factors, and the potential bias stimulus familiarity. It would appear from the results of analysis that the constructs advanced by the model possess an adequate degree of independence. The testing of the constructs are not limited to the specific directives used in the present prototype instrument. Certainly other directives could be used in the scenarios to assess different component skills. In fact, other component skills were contemplated in the development of the persent instrument. These happened to be some of the most frequently identified component skills from previous studies (Peterson, 1982, Peterson and Watkins, 1978). In this regard, the author is not satisfied with the Valuing directive. The directive used in this battery represents a skill too far down in a hierarchial analysis of the Valuing dimension. Nevertheless this particular directive pertaining to the infering of values from behavior is almost universal in the lists of generic skills proposed by liberal arts faculties.

These generic skills do cut across assessment tasks. A recent study by Sackett and Dreher (1982) found that the mean correlation among typical assessment center tasks (In basket, Leaderless group, etc.) was near zero,

raising questions concerning whether the concept of generic skills is relevant. However, the variability in the kinds of exercises in the present prototype measure is much less than in the typical "bill-of-faire" in assessment center tasks. Nevertheless, an issue is raised for further research, regarding limits of generic skills in terms of decreasing task fidelity or epistomological distance. Had one of the tasks required mathematical proficiency, would the intercorrelations among exercises be reduced? The fact that the average attenuated correlation different among skills (r_{tt}=.69) was greater than the average attenuated correlation between the same skills across tasks (r_{tt}=.49) gives support for further inquiry into the meaning of the concept generic.

Another interesting finding was the extent to which the skills co-vary with educational experience and age. Cattell (1977) and Horn (1981) propose that general ability consists of two abilities: Fluid and Crystallized abilities. Fluid abilities, often measured by figural analogies, classifications, series completion problems, memory span and the like, begin their development early in life and reach a peak in late adolescence and then begin to decline in early to middle adulthood. Crystallized abilities on the other hand, often measured by tests of vocabulary, reading comprehension, semantic relations and general information, tend to increase indefinitely with age. These latter abilities appear to be similar to the ways in which generic skills, as measured in this battery, perform. It seems logical that educational experience would have an impact on the development of crystalized abilities but not on fluid abilities. Nevertheless, the relationship between generic skills and fluid abilities is point of

inquiry. The use of the Wide Range Vocabulary Test was an attempt to accountfor a fundamental attribute (i.e. a fluid ability) of higher order intellectual performance but it may actually be another construct in an array of crystallized abilities as suggested by Willis, Schail and Lueers (1983). Nevertheless, when using the battery to compare the development generic skills across academic disciplines or between different classes, some measure should seemingly be used to equate potential differences in academic aptitudes. The Wide Range Vocabulary Test would ostensibly be an indicator of verbal aptitude.

Irresistably, one last "as if" analysis was conducted to determine whether there would be significant differences between lower division, upper division and graduate students in this sample population in terms of generic skill performance. In no way should these results be construed as representing these respective groups at Florida State University. The subjects were not drawn at random from circumscribed populations. Nevertheless, data are presented to demonstrate how the measure might be used to evaluate the growth of generic problem-solving skills during the baccalaureate years. Verbal ability as measured by the ETS Wide Range Vocabulary Test was used as a covariate to equate verbal ability among these educational levels.

The results of this "demonstration" analysis are portrayed in Table 6 on the next page. One can note there were no differences between lower and upper division performances but that graduate student performances across five or six skills were significantly different from undergraduates. Could it be that undergraduate education enhances the development of cognitive schemas on which to structure and interpret subsequent life experience?

Table 6

Generic Skills by Performance by Educational Level with Vocabulary Score as Covariate.

Skill		Lower Division (N = 20)	Upper Division (N = 26)	Graduate (N =16)	F
Proble Solving	M	22.6[a]	21.8	30.9	7.97***
	SD	5.0	4.8	6.1	
Communication	M	23.8	23.2	28.9	.88
	SD	5.2	8.6	6.4	
Analysis	M	24.1	23.9	31.3	4.74*
	SD	3.1	4.7	4.4	
Synthesis	M	23.5	23.5	32.1	6.18**
	SD	3.4	5.3	5.9	
Valuing	M	15.6	17.3	24.9	4.80*
	SD	3.3	5.1	7.4	
Execution	M	22.0	21.9	30.5	8.39***
	SD	4.5	4.5	5.6	

a. Unadjusted Means

* $p < .05$
** $p < .01$
*** $p < .001$

Perhaps a requisite amount of life's experiences is required before higher-order thought processes become evident in the real-life level problem situations. It could be that problem scenarios cast at the major field level would discriminate between lower and and upper division students.

Reliability. The battery appears to possess sufficient reliability for use as a program evaluation instrument. The modest interrater product-moment correlations could be improved by a more concentrated effort to train raters and to provide an honorarium to foster motivation to make subtle discriminations among performances. With such procedures, the assessment battery could be reduced to two exercises and rated by two raters without appreciably lowering reliability. The cost of administering and scoring would also be reduced.

Utility. Assuming that the instrument is sufficiently valid and reliable for use in academic program evaluation, administrators ultimately inquire about the costs to the institution, the faculty and to students. The estimated direct costs to develop, administer, score and analyze the results of this instrument for a subject pool of 62 was $6000. This amount does not include research costs involved prior to the actual development of the measure. It must be remembered however that costs escalate with the desire to achieve greater reliability in scores and greater confidence that a sample respresents a given population. In order to obtain samples of freshmen and seniors that possess a sampling error of less than .05 and a confidence level of .10, as many as 400 subjects would be required (Cochran, 1963). At this confidence level, the cost of evaluation escalates to $15,000 assuming 3 exercises, 3 raters, and paying each subject $10 to achieve a desired show-rate for stratified random sampling procedures.

With two exercises and two ratings for each exercise the cost is $11,000. The major variable costs, of course, are scoring and paying students to take the test to achieve a desired sampling accuracy. Costs of course are reduced as these two factors are compromised. Costs for instrument development, computer analysis, clerical support and supplies are nearly fixed and these were estimated to be $4500.

Conclusion. Without a doubt, performance testing is expensive compared to using commercially prepared multiple choice tests. The usefulness of performance tests will depend on whether the kinds of information about students are useful for making curriculum decisions pertaining to course structure and student degree requirements. A previous study (Peterson, 1982) reported that when performance tests and multiple choice tests are combined in an assessment battery, these two kinds of assessments load on two distinct factors suggesting that performance tests do provide different information from multiple choice tests. A prototype baccalaureate measure has been developed that appears to be associated with the level of educational attainment (i.e. total credit hours completed and level of educational attainment). However, the relative contributions of maturation, certain intellectual factors and various educational experiences to generic skill performance should be ascertained through experimental research paradigms before curriculum decisions should be based on its use. Nevertheless, this prototype measure appears to be worthy of further exploration as a research instrument with which to understand more about the relationship between curriculum and outcomes of postsecondary education.

REFERENCES

Astin, A. W. Lets try a "value added" approach to testing. Chronical of Higher Education, July 28, 1982.

Branson, R. Task analysis procedures extended to non-task behaviors. Paper presented at Annual American Educational Research Association meeting, Boston, 7 - 11 April 1980.

Cattell, R. B. Abilities: Their structure, growth and action. Boston: Houghton-Miffin, 1971.

Cockran, W. G. Sampling techniques. New York: Viley, 1963.

Flavel, J. H. The developmental psychology of Jean Piaget. New York: D. Van Nostrant Co., 1963.

Forrest, A. and Steele, J. Summary report of research and development: 1976 - 1980, College Outcome Measures Project. American College Testing Program, Iowa City, Iowa (In Print).

Goleman, D. The new competency tests: Matching the right people to the right jobs. Psychology Today, 1981, January, 35-46.

Horn, J. L. Organization of abilities and the development of intelligence. Psychological Review, 1968, 75, 242-259.

Katz, J. It's time to reverse our retreats from reality in teaching college students. Chronicle of Higher Education, July 21, 1982.

Kohlberg, L. Stages of moral development. In C. M. Beck, B. S. Drittendon and E. V. Sullivan (Eds.) Moral Education, Toronto: University of Toronto Press, 1971.

Lewis-Beck, M. S. Applied regression: An introduction. Beverly Hills, Ca.: Sage University Papers, 1980.

McClelland, D. C. Testing for competence rather than for intelligence. American Psychologist, 1973, 28, 1-14.

Perry, W. G., Jr., Forms of intelligent and ethical development in the college years. New York: Holt, Reinhart and Winston, 1970.

Peterson. G. W. A meta-evaluation of a generic skills approach to evaluating academic programs. Paper presented at Annual Convention of the American Research Association, New York, 23 March, 1982.

Peterson, G. W. and Stakenas, R. G. Performance-based education: A method for preserving quality, equal opportunity and economy in public higher education. Journal of Higher Education, 1981, 52, 352-368.

Peterson, G. W. Competency-based articulation: Clearning up the confusion in community college-university transfer. Improving College and University Teaching, 1981, 29, 169-173.

Peterson, G. W. and Ramsey, M. G. A methodology for measuring officer job competence. Paper presented at American Psychological Association Annual Meeting. Divison 19, Los Angeles, August 1981.

Peterson, G. W. and Watkins, K. Identification and assessment of competence: Final report. Florida Competency-based Articulation Project. Resources in Education (ERIC), ED 169839. September 1978.

Sackett, P. R. and Dreher, G. F. Constructs and assessment center dimensions: Some troubling empirical findings. Journal of Applied Psychology, 1982, 67(4), 401-410.

Warren, J. R. ETS Measures of General Education. Berkeley, Ca: Educational Testing Service, 1980.

Willis, S. L. Schail, K. W., and Lueers, N. Fluid-crystallized ability correlates of real life tasks. Paper presented at annual American Educational Research Association Meeting in Montreal, CA., April 1983.

Winter, D. G., McClelland, D. C., and Stewart, A. J. A new case for the liberal arts. San Francisco: Jossey-Bass, 1981.

Woditsh, G. A. Developing generic skills: a model for competency-based general education. CUE Project, Bowling Green State University, Bowling Green, Ohio, May 1977.

_____, Wide Range Vocabulary Test, New Jersey: Educational Testing Service, 1962.

APPENDIX I

Real Life Problem Solving Exercise
"Cuban Crisis"

PART I: The Decision

The Scenario

As a case worker for the Department of Social Services, you have been assigned to work with a young Cuban woman, found beaten and semi-conscious in a Cuban neighborhood. Taken by police to a hospital, she acknowledged that she had been sexually assaulted, but refused to identify her attacker; she became hysterical at the suggestion that she undergo a physical examination to determine if rape had occurred. Her family has refused to cooperate with police attempts to investigate the presumed rape, although the police have been able to determine her attacker's identity by informants in the Cuban community. The girl's father has flatly refused to swear out a warrant, and the girl refused to acknowledge either that a rape has occurred or her assailant's identity. Frustrated, the police have called you to come to the police station and take over the case.

Directions:

Your task is to describe which course(s) of action you would recommend be taken and why. Please elaborate as much as you can. Use the back of this page if necessary.

Part II: The Inquiry

The Scenario

As a case worker for the Department of Social Services, you have been assigned to work with a young Cuban woman, found beaten and semi-conscious in a Cuban neighborhood in Miami, Fla. Taken by police to a hospital, she acknowledged that she had been sexually assaulted, but refused to identify her attacker; she became hysterical at the suggestion that she undergo a physical examination to determine if rape had occurred. Her family has refused to cooperate with police attempts to investigate the presumed rape, although the police have been able to determine her attacker's identity by informants in the Cuban community. The girl's father has flatly refused to swear out a warrant, and the girl refused to acknowledge either that a rape has occurred or her assailant's identity. Frustrated, the police have called you to come to the police station and take over the case.

Directions:

Please answer the following questions as completely and concisely as you can. However, if there is not enough space, please ask the proctor for additional paper.

Communication

1. As a social worker, what are the central issues for you in the above scenario?

Analysis

2. Describe the problem from the girl's perspective, from the father's perspective, from the police's perspective.

(a) The girl's:

(b) The father's:

(c) The police:

Synthesis

3. List as many possible courses of action that might be taken in the above scenario, including options you would not choose to follow. Use the back of this page if there is not enought space for all of your alternatives.

(a)

(b)

(c)

(d)

(e)

Valuing

4. What values might you hypothesize underlie the father's behavior? The girl's behavior?

Father's values:

Girl's values:

Execution

5. For the solution you recommended on page 1, outline the sequence of actions that you would take to implement the solution. Describe what you would do first, second, third, etc. (Go on to the back of this page for additional steps).

(1)

(2)

(3)

(4)

Familiarity Scale

Indicate your familiarity with Latin or Cuban culture by circling the number which most closely approximates your acquaintance.

1. know nothing (e.g., you have no knowledge of Cuban or Latin culture nor how these cultures are different from Middle America.

2. only a vague familiarity (e.g., your familiarity is restricted to what you have read in the newspapers or have seen on television news broadcasts)

3. somewhat familiar (e.g., you have studied the culture but do not speak or read Spanish)

4. familiar (e.g., you have studied the culture and have taken courses including Spanish)

5. very familiar (e.g., you have first hand acquaintance through living in the culture for 6 months or more and speak the language. You are very familiar with ways in which the culture is different from Middle America.)

Generic Skills Rating Scales
Cuban Crisis

I. Generalized Problem Solving (Formulate a plan of action to rectify a need or anomaly based on a higher order, principled rationale)

High 5	4	Medium 3	2	Low 1
Options selected and reasons given are logical and demonstrate capacity for higher order principled thought. The major concern to others should be the girl's psychological well-being. The proposed action is concrete (as opposed to vague, abstract) and several perspectives are brough to bear on the decision There is a consideration of social and cultural factors.		Solution attempts to satisfy everyone's needs such that no one's needs are completely met. Or, girl's psychological well being sacrificed for apprehension of rapist. However, their must be some concern for girl's psychological well-being. At this level, there is no indication of a balance among perspective. Focus is only on one part of scenario, i.e. the girl's welfare or bringing the assailant to trial.		Solution is vague, ambiguous, trial and error strategies used. Solution does not take into account multiple perspectives and values operating in the drama. Favors one view with out justification. Solution demonstrates no regard for girl's well-being, eg. "Get the rapist".

II. Communication (Understands of complex situations by expressing the main idea in ones own words.)

High 5	4	Medium 3	2	Low 1
Mentions that the first priority is the girl's psychological well being. Answer reflects sensitivity to Cuban cultural issues ("machismo", "chaperoning", etc). Mentions father's need to support and answer addresses society's responsibilities to individuals in need. Sensitivity to "deeper", "latent issues".		First emphasis is on capturing and punishing rapist. Mentions concern for girl's well being, cultural issues, fathers needs for support, but these are secondary. Answer reflects incomplete or superficial awareness of underlying issues eg. "Help the girl and family feel better".		Responds to manifest issues, eg. "getting all parties to cooperate with police". No concern for girl or father. No awareness of cultural issues, intolerant, prejudiced biased. Extremely superficial understanding of the complexities of the situation. Completely neglects the girl.

III. Analysis (Perceive elements of problem situation as evidenced by the ability to view an event from multiple perspectives).

High 5	4	Medium 3	2	Low 1
From the girl's perspective: denial of rape to protect family or personal honor; obedience to the father. Father's perspective: guilt about not living up to cultural norms regarding protection of daughter; shame regarding daughter; wants to manage crisis privately. Police's perspective: social control need to see justice served. All three perspectives recognized and plausibly correct.		Answer reflects awareness of only two perspectives or one perspective emphasized more than the others. Simple analysis, but "on the right track". May be vagueness about one of the perspectives. Some understanding of cultural issues.		Only one perspective recognized. Answer based on information not present. Analysis very superficial, opinionated, biased prejudiced, lacks any understanding of cultural issues. Projects beyond information given (eg. "she fears she's pregnant and doesn't want to exam, or "knows the rapist and is afraid to tell" or "fears additional attacks".

IV. Synthesis (Formulate a variety of discretely plausible courses of action)

High 5	4	Medium 3	2	Low 1
At least 3 to 5 plausible and discrete courses of action outlined. Each clearly presented without bias. Answer reflects that the option selected in PART I was not the only possible one.		Three to five options presented but with an obvious bias toward one of the options. Does not reflect a "freeing up" of the mind to entertain other alternatives. Imbalance in the clarity of options. Options are not discrete, but are restatements of two or three options.		Only one real option listed. Option implausible or unclear. Question not answered (e.g. gives strategy for how one course of action might be implemented). Illogical "Put attacker, girl and father in one room and see what happens".

V. Valuing 1. (Infer values from behavior)

High 5	4	Medium 3	2	Low 1
Three plus two values identified for each role (5 total). Girls: obedience, avoidance of conflict, chasity. Fathers: control, privacy, family honor. Answer reflects familiarity with cultural and societal values.		Most obvious values identified (3) but answer reflects lack of understanding of cultural and societal values brought to bear in this situation. Emphasizes only the father's or daughter's values.		Little or no awareness of what values are, e.g. "The girl did what she had to do", or "she feels sorry for herself". Mistakes feelings for values. Expresses motives for values, eg "doesn't want to get involved".

VI. Execution (Perceive logical means-ends relationships with sufficient detail)

High 5	4	Medium 3	2	Low 1
Description is such that someone else could carry out action. There is a logical and clear progression of steps - a "flow chart" in words. Optional actions stated, sequential, and lead to closure. Aware of preparation before interviewing.		The description is general; the sequence of actions is broken; important steps missing; some irrelevant information. Individual could "muddle through" to the solution.		Description not clear, vague, over simplified. Someone would not be able to implement solution from the description of events. Trial and error approach to implementation. Answer does not lead to a solution-only an array of activities.

APPENDIX III

The Scenarios

Cuban Crisis (I)

As a case worker for the Department of Social Services, you have been assigned to work with a young Cuban woman, found beaten and semi-conscious in a Cuban neighborhood. Taken by police to a hospital, she acknowledged that she had been sexually assaulted, but refused to identify her attacker; she became hysterical at the suggestion that she undergo a physical examination to determine if rape had occurred. Her family has refused to cooperate with police attempts to investigate the presumed rape, although the police have been able to determine her attacker's identity by informants in the Cuban community. The girl's father has flatly refused to swear out a warrant, and the girl refused to acknowledge either that a rape has occurred or her assailant's identity. Frustrated, the police have called you to come to the police station and take over the case.

Land Development (II)

You are a County Commissioner and a special meeting has been called in order to settle a dispute that has occurred concerning the proposed development of 100 acres of land along a primitive river bank. The land in question is 25 feet beyond a marshy area that borders the river. The developer plans to subdivide the land into one acre residential tracts. The site is in a rural county that lies just 7 miles south of a rapidly growing metropolitan area. The river and surrounding areas are noted for their excellent fishing and camping, annually attracting sportsmen and campers statewide and beyond. A major controversy has been raging ever since the developers intentions were made public. Local environmental groups here opposed the development while local businessmen have stressed the need for such a development. The developer had promised to delay the actual construction until environmental groups completed an environmental impact study. The special meeting has been called because several days ago the developer began construction of a road along the river bank and has also begun digging wells and septic tanks. The local environmental groups are demanding an immediate halt to construction. The developer claims that the construction delay is costing him thousands daily and that he intends to proceed with construction until concrete evidence is provided that the subdivision would be other than beneficial to the community.

Token Economy (III)

You are a School Board member in an inner city school district. Disciplinary problems in the 3rd through 6th grades have tripled in the past year. The School Board has received numerous demands from parents to do something to restore order to the schoolrooms. Because of understaffing the classrooms are overcrowded, and it has become increasingly more difficult for teachers to manage the students. Fighting, practical joking, inattentiveness, the use of vulgar language and cheating abound. Several principals have asked the School Board to institute a token economy in all elementary school classrooms within the school district. Since this proposed solution was introduced at the last board meeting a controversy has raged between opposing factions in the P.T.O. Those favoring the institution of a token economy claim it is the most efficient way to eliminate undesirable classroom behavior. Those opposing the use of a token economy do so primarily on ethical and moral grounds. The board meets next month to make recommendations for further courses of action.