**ASSESSMENT OF MINISTRY PREPARATION TO INCREASE UNDERSTANDING**

John Harris
Samford University

Dan Aleshire asked me to discuss assessment strategies and techniques for seminary education with him and his staff. As the day was ending, he asked me to outline actual assessments for a seminary. His question forced me to cut through abstractions and theory to put something specific and operational on the table.

Aleshire's penetrating question stayed with me and I decided to put something on paper. In October Aleshire came to Samford to discuss a preliminary draft of this paper with Brad Creed, currently provost at Samford and formerly the founding dean of Baylor's Truett Divinity School; Linda Cannell, professor of Educational Ministries at Trinity Evangelical Divinity School and member of the ATS Executive Committee; and me. While these meetings added to my understanding of the challenges of assessment for theological education, Aleshire, Creed, and Cannell are not responsible for the thoughts that follow.

One may well ask, "What does John Harris bring to a discussion of assessment of theological education?" The first clear answer is, "Not enough." I have spent most of my 30-plus years in higher education working on assessment and accreditation issues along with efforts to improve teaching and learning. My acquaintance with theological education includes an undergraduate major in Bible and two years of seminary before deciding to go into education, teaching M.Div. and D.Min. leadership courses and serving on seminary accreditation teams. Having no substantive understanding of Roman Catholic seminary formation of priests, I recently spent five days living among the seminarians and faculty of Mundelein Seminary. This brief experience certainly makes me no expert on priestly formation, but it did open my eyes to how formation is done in community as compared to the course and classroom emphasis of Protestant and Evangelical seminaries. Nevertheless, my thoughts about assessment of theological education inevitably reflect my greater knowledge and experience with Protestant and Evangelical theological education.

<u>Purpose-driven Assessment</u>. This paper focuses on assessment of ministry preparation as distinct from graduate theological education. For some, these concepts may be the same. From my perspective, that would be like saying that a Ph.D. in biochemistry or in other medical sciences is the same as an M.D. in education *and* training.

I am reminded of T. Hale Ham's description of grand rounds of medical education's history. Ham, a hematologist on the faculty of Case-Western Reserve Medical School, pioneered medical students learning the science and practice of medicine concurrently rather than two years of pure science followed by clinical training. According to Ham, the first round of medical education was apprenticeship to a doctor, which left much to be desired in knowledge of the medical sciences. The second round emphasized medical science, with less emphasis on learning to become a physician. The third round integrated medical science and practice.

In the 1970's, Ham led Case-Western Medical School to put first-year medical students in groups that learned about the cell and participated in prenatal care and delivery. In the fourth year, students studied geriatrics and cared for elderly patients. Medical education has led higher education in active learning strategies such as problem-based learning and the use of trained actors as "mystery patients." The ultimate question is not whether physicians can answer examination questions about disease characteristics, but whether they can use their medical knowledge for accurate diagnoses.

The grand rounds of seminary education seem to have followed a somewhat similar path in its first two rounds but has yet to embrace the third fully. In the first round, seminary faculties were predominantly comprised of former ministers. In the second round, Ph.D. scholars with limited experience in full-time ministry became the mainstay of seminary faculties. A third round to integrate scholarship and practice has not occurred generally. Protestant and Evangelical seminary education continues to be a classroom-based academic experience led by Biblical and theological scholars. The disconnect between fresh Ph.D.s in their first year of teaching is such that ATS provides a workshop to orient them to seminary education.

<u>Appearance or Value</u>. The apparent disconnect between the training and interests of seminary faculty and ministerial preparation is important for any discussion of assessment of ministry preparation programs. If assessment is to be worth its cost, it must be derived from and congruent with the actual and not just rhetorical purposes of an educational program. The assessment strategies proposed in this paper are for ministerial preparation, not for graduate, research-oriented, scholarly theological education. Assessments for the latter might include externally reviewed theses, comprehensive examinations by multiple readers, and success of graduates in relevant Ph.D. programs. However, I believe assessments for ministry preparation can and should include such assessments; but as suggested below, these other types of assessment are also needed.

Given the pressure for assessment of effectiveness from the public in general and accrediting agencies in particular, administrators and faculty often assume that they can satisfy those demands by tacking assessment onto traditionally formatted curricula, conventional instructional practices and discipline-based organization of the school. That is, surveys of students and alumni, and on rare occasions, comprehensive assessments of learning beyond tests within courses, are added to existing programs, often just a couple of years before a re-accreditation review is due. As a member of a number of accreditation visiting teams, I usually am assigned to review assessment programs. More frequently than not, I have observed hastily constructed, Potemkin-village assessment programs. I often sensed that these programs would be quickly collapsed after the visit. To the typical professor and many administrators, it all seems like busywork and much-ado about nothing; sadly enough, that is exactly what it is in too many cases.

Assessment for Optimization. There is really no substantive, lasting value in assessment unless faculty and administration are clear about their actual purpose and are determined to deploy themselves and their resources to fulfill that purpose. If a seminary determines to optimize its effectiveness in preparing ministers, it will need the following assessments:

- Screening and selecting applicants for abilities, habits, experiences, values, and attitudes that predispose them to benefit from what the seminary offers, and that bode well for perseverance and fulfillment in ministry in the principal employing churches or other ministries.

- Determining characteristics of faculty that make them most effective in preparing ministers, and using those characteristics as criteria in selecting, developing, and retaining faculty.

- Determining and evaluating curricular content, learning strategies, and communal life that most effectively shape individuals for a lifetime of ministry.

- Tracking the academic, spiritual, and personal development of each seminarian from admission to graduation.

- Studying what happens intellectually, spiritually, and personally to students as they move through a program.

- Tracking the careers of graduates to identify the characteristics that distinguish those who persist and find fulfillment in ministry, and those who do not.

The key to optimizing any organization is to increase corporate understanding of how it works. Peter Senge's well-known *The Learning Organization* provides ways of corporate learning that with some translation can help a seminary faculty increase its understanding of its work. But as a former executive of a major city's electrical utility told me, "I go to management training sessions and come home to find that the best concepts are in the New Testament." Assessment of minister preparation programs can and should make use of psychometric theory and instruments, but it must also be theologically informed and shaped. It seems that theological educators have sometimes adopted social science strategies and ways of knowing without weighing the theological ramifications. The commonly employed educational assessment strategies rest on positivist assumptions. Numbers are useful, but the more one knows about how they are collected and analyzed, the more one realizes that they serve best as indicators of matters that should be probed more deeply through tacit knowing. Dennis Sansom and I, in *Discerning Is More Than Counting,* attempted to describe the assumptions on which purely quantitative assessment strategies rest and their consequent limitations.[i]

**Assessment to Increase Understanding**

When one moves from generalities and theory to proposing specific assessment, disagreements intensify. In all the years that I wrestled with assessment, I do not recall a single strategy or technique about which honest and valid concerns could not be raised. Those most familiar with the phenomena of being assessed sense that the measure does not fully reflect its reality, which raises the epistemological question of how completely any anything be known or comprehended, much less fully measured. I do not propose these strategies and techniques as complete measures of the human qualities in question. Objective and quantitative assessments, particularly of human behavior, serve only as "leaves in the wind." They are only indicators useful for prompting further reflection by multiple interpreters.

Therefore, I propose assessments at three points--admission, during the program and after graduation--to stimulate and focus ongoing, *community* reflection on ministry preparation. The purpose of this ongoing, communal reflection is to increase understanding of the inter-relationships among 1) the types of students admitted; 2) the mix of curricula, instructional methods, and faculty characteristics that comprise the seminary program; and 3) the perseverance and performance of graduates in ministry.

**I. Admission Assessment**

Assessment of applicants and newly matriculated students can serve three important purposes. First, assessment may be used to identify individuals with undesirable backgrounds, behavior/personality disorders, or obvious lack of preparation for study and formation. Second, assessment may be used to select applicants most likely to succeed and persevere in seminary and ministry. Third, once an entering class is established, assessment may be used to take stock of the students admitted.

Assessment for Screening. Seminaries may use some combination of criminal background checks, credit reports, health reports, documentation of vaccinations, personality tests, and psychiatric interviews to identify individuals who would not function well within the seminary community. A seminary also might decide to screen applicants for theological compatibility and denominational identity. That is, a seminary might decide there is a range of beliefs and behaviors within which it can work with students; consequently, it might use questionnaires or interviews to identify individuals most likely to profit from its instruction and community.

Timothy George, dean, Beeson Divinity School, Samford University, suggests that churches could contribute to effective screening of applicants.[ii] An applicant's home church will usually know the applicant's typical behavior patterns. I am impressed with how thoroughly Roman Catholic diocesan vocational directors review candidates for admission to seminary. There are major differences between Roman Catholic and Protestant and Evangelical polities which produce the differences in the church endorsement approaches. Nevertheless, George suggests that Evangelical churches could and should do more than they often do reviewing an individual asking the church to endorse their application for admission to seminary. A brief chat after a Wednesday evening prayer is not enough.

Given the absence of uniform policies and processes in the ”free church” tradition, seminaries will likely have to take the lead in developing and promoting more systematic review processes leading to church endorsement. Such a review process should include relevant discussions about beliefs and commitments; but to add maximum value, it should also include a kind of observed life-history of the applicant. Obviously, the applicant’s participation in the life of the church could be documented, but also the applicant’s school and work history should be included. Perhaps no component of one’s life-history is more important that their family narrative. While we know that patterns of past behavior are the best predictors of future behavior, all such predictions are based on low empirical correlations. Religious history is full of individuals who by their efforts or divine intervention rose above their pasts. So the review should not only include a report of influential circumstances and prevailing behavior patterns, but also life-changing religious experiences. And in the end, an admissions committee will remember prophets are seldom welcomed in their home towns.

Assessment for Selection. As most educators know, educational outcomes are greatly determined by the quality and nature of admitted students. In fact, there are empirical studies that demonstrate that the effects of educational treatments or curricula on intellectual and knowledge achievements of graduating students are minimal when incoming intellectual differences are statistically controlled. In other words, incoming abilities explain more of the variation in graduates’ academic achievement than is explained by the educational program.

Ability tests and grades only moderately predict future grades; for example, tests such as the GRE and undergraduate grades at best will account only for about one third of the variance in seminary grades. Because of their lower reliability, psychological assessments and interest inventories are generally even poorer predictors of future behavior. We should also be aware that judgments based on interviews correlate poorly with later performance. This is true even when the interviewers are well trained, such as clinical psychologists. If human behavior is completely determined by genes and environment, as some believe, experts have yet to identify the variables and means of measuring that predict individual human behavior at the accuracy level of physical sciences.

While precise prediction of individual behavior is not available, there are techniques and instruments that predict certain performances significantly better than chance. For example, the modest correlation of previous grades and test scores with future grades allows one to predict rather accurately the extremes of a distribution. That is, the farther an applicant’s previous grades and test scores are below the average of all the applicants’ grades and test scores, the greater the chance the applicant’s future grades will be at the lower end of the class’ grade distribution. The opposite is equally true. This means an admission committee is on fairly safe ground in admitting the applicants at the top of the distribution of applicants’ grades and test scores as far as their seminary academic performance is concerned. The chances of the applicants succeeding academically who are far down in the distribution of entering grades and test scores are small. But modest correlations do not help much in judgments about applicants' grades and test scores in the middle range of the distribution.

Whatever emphasis a seminary may place on academic achievement, one assumes it also will be interested in the student's personal and spiritual development. Personal and spiritual development may have more to do with perseverance and fulfillment in ministry than academic performance. Yet, as messy as assessment and prediction of academic performance are, assessment and prediction of attitudes and behaviors are even messier. Measures of cognitive abilities and academic achievement require the test taker to demonstrate knowledge, skill, or understanding. In contrast, when one responds to an inventory or questionnaire about interests, values, or emotional states, one struggles with several problems, such as what should I report versus what is expected or socially acceptable. That is, questionnaires and inventories about attitudes, interests, and behaviors as one to report typical or usual feelings and behaviors as distinct from what they might perceive to be expected or desired. While tests of cognitive abilities can be affected by circumstances in which the test is administered, and the person's physical and emotional condition, personality inventories may be more affected by them.

Cronbach[iii] made a useful distinction between tests of ability or maximum performance and tests of typical performance. Ability tests for admission, such as the GRE or Miller Analogies Test, assess maximum performance. One is expected to do one's best. Questionnaires, such as the popular Myers-Briggs Type Indicator or the heavily researched Minnesota Multiphasic Personality Inventory and Strong-Campbell Interest Inventory, ask the respondent to report typical or usual behaviors. As a genre, such inventories or questionnaires are less reliable or consistent than maximum performance tests of abilities or skills. Consequently, their correlation with academic or work performance is lower.

Bio-Data Prediction. Given that so much is a stake in assessment for screening and selection, seminary educators might find value in supplementing ability tests and typical performance questionnaires with bio-data prediction. In brief, bio-data prediction involves in-depth analysis of individuals' biographies to identify factors that predict future performances or behaviors. That is, the researcher is working with what the individual or others report about what they have actually done, i.e. service activities, church work, extracurricular school activities, hobbies, etc. The following description of bio-data prediction may be helpful in understanding what is involved.

> Biographical inventory is a selection device used as an alternative or supplement to cognitive testing because this measurement method predicts aspects of job performance that are not predicted by cognitive measures. Examples of these aspects of performance are interpersonal relationships (e.g., with co-workers or clients) and motivation.

Biographical inventories have been empirically developed against such varied criteria as amount of insurance sold by life insurance agents, turnover of bank clerks, productivity of research scientists, and performance of naval personnel in diver training. Such inventories have proved valid as predictors of job performance in groups ranging from unskilled workers, office clerks, and service station dealers to chemists, engineers, and high-level executives. Personal history types of items that discriminate can provide a great deal of information about what kinds of employees remain on a job and what kinds do not, and what kinds are promotable and what kinds are not.

RATIONALE
The assumption that underlies the use of bio-data is that past behavior is a valid predictor of future behavior. More specifically, it is assumed that information obtained from job applicants about previous work experiences, education, etc. can be used to predict job performance. Items included in these inventories are selected on the basis of previous research which demonstrates significant relationships between item responses and job performance. Personal history items commonly used fall into the following areas:
- personal
- background, general
- education
- employment experience
- skills
- socioeconomic level-financial status
- social
- interests
- personal characteristics, attitudes expressed[iv]

In addition to providing an admissions committee information on variables related to success in seminary, individuals could their bio-data reports to reflect on their call to ministry. That is, after completing a structured bio-report, one considering ministry might reflect the bio-data to discern paths and patterns that confirm or question one's call to ministry. That is, one might review the biography for transition points that seem to have moved one toward ministry. Obviously, we are prone to report and see in what we reported what we want to see. So sharing one's review with a mentor or counselor would be desirable on most cases. An admissions committee, in addition to reviewing bio-data reports for factors that are empirically related to success in seminary and ministry, could ask candidates assess their call to ministry by the paths and patterns that observe in their bio-reports.

<u>Assessment for Optimal Learning</u>. More is at stake than deciding whom to admit. Those preparing others for ministry need to understand their students. Ideally, they should have some shared understandings or intersubjective agreements of the desired personal traits, ministry skills, and theological abilities that beginning ministers should possess. Given these understandings and agreements, they may determine a pattern of the differences between the entering students' traits, skills, and abilities, and those they

hope to nurture in students by the completion of the seminary experience. Once a beginning class is admitted, the faculty should develop some empirical and tacit understanding of the material with which it has to work. That is, how do the entering, first-year students look in comparison to what it is hoped they look will look like three or four years later? The gaps between what they enter with and what they should leave with defines the seminary's education or formation task. Only when education or formation is seen as closing the gaps does assessment add value equal to its costs.

## II. <u>During-Program Assessment</u>

ATS expects member institutions to assure four outcomes in M.Div. programs:

a. Knowledge of religious heritage
b. Understanding of cultural heritage
c. Growth in spiritual depth and moral integrity
d. Capacity for ministerial and public leadership

In light of these expected outcomes, a comprehensive, systematic assessment program should be developed to provide feedback on the program's effectiveness in facilitating student achievement of in each of these four goals or desired results. Ideally, such assessments should be embedded in the total curricular experience, classroom and community life. When a goal involves cognitive learning, the assessment should test for understanding and the ability to act on the learning, such as applying the learning to an issue or problem different from the context of the original learning. These four goals clearly require integrative understandings and behaviors that cross theological disciplines and the confines of modular courses; therefore, the assessments should be interdisciplinary and require integrative responses. The four goals also contain mixtures of cognitive and affective-behavioral learning. For example, knowledge of religious heritage appears to call for a straightforward test of knowledge, whether at the recall level or at higher cognitive levels, such as application, synthesis, or evaluation. "Growth in spiritual depth and moral integrity" obviously calls for a very different type of assessment. In this case, one is interested in typical attitudes and behaviors that are more amenable to assessment through surveys, self-reports, and unobtrusive observations.

Examples of assessment strategies and techniques for each of the four ATS goals are offered below.

<u>Knowledge of Religious Heritage</u>

*Annual Theological Problem*. Present each year's cohort (First-year, second-year, and third-year cohorts) with a challenging theological problem. The aim is not to determine if students reach certain predetermined theological conclusions, but to assess their abilities to think theologically. There are many techniques of juried assessments, e.g., creating divider essays between three categories–outstanding, acceptable, and unacceptable. Each paper could be read by two readers independently; if both rate it the same, that is the rating it would receive. If the two disagree, a third reader would rate the paper. Once the papers are sorted into the three categories, the same or other readers could compare and contrast the outstanding and unacceptable papers to identify patterns of differences between them. Another strategy would be to compare the students' papers with three or more papers on the same issue by professors. Yet another strategy to determine improvements in a given year is to have the students write on the same issue twice in one year–one paper at the first of the fall and the second at the end of spring. Consistent improvements would suggest what the curriculum actually yielded in terms of student learning that year.

*Charlemagne-Alcuin Encounter*. Perhaps the best test of a pastor's ability to think theologically is the ability to respond to the person in the next on an airplane who asks a question about common but complex issues, such as the Trinity, predestination-free will, or the resurrection. Though illiterate, Charlemagne asked Alcuin, the theological scholar, questions that rattled his brain. Similarly, M.Div. students could be tested by the questions that theologically unschooled but bright laypeople, adults and children ask. Such questions would challenge the student's integrated understanding of theology and the student's ability to explain complicated theological issues in everyday language. Ideally, students would be quizzed by trained actors.

*Theology and Classical Issues*. M.Div. students could be presented with recurring dilemmas in human existence through the classics of great cultures, such as Confucius, Homer, Plato, etc. The test would be twofold: 1) demonstration that they understand the perennial human problem as presented by the classical figure and the person's proposed resolution of it; and 2) their ability to deal with the problem from a Christian perspective.

<u>Understanding Cultural Context</u>

*Reading a Different Socioeconomic Setting*. Have M.Div. students spend several days in a community very different economically and socially from their own. Ask students to identify the primary recurring values, power structure, and economic patterns of the new culture. Against the backdrop of their analysis, ask them to formulate the most effective expression of Christianity and ministry for that culture.

*Reading One's Own Setting*. Ask students to describe their own culture using basic anthropological, sociological, and economic tools. Ask them to describe how that culture has shaped its current expression of Christianity and to describe how the church could be a more faithful witness to that culture.

*Reading the Effect of a Historical Moment*. Choose a particular time and place in Christian history, and ask students to identify and describe how that time and place shaped Christianity and the church as an institution. Then ask the student how well or poorly that particular culture shaped Christianity and how the church witnessed to its culture.

*Connecting Multiculturally*. Arrange for every student to work in a mission setting in a culture distinctively different from the student's own culture. Ask the resident missionary to observe and report how well the student comes to understand and develops rapport with the people.

Moral and Spiritual Growth

As I approached suggesting assessments for moral and spiritual growth, I was reminded of an oxymoronic session at an annual Assessment Conferences of the American Association of Higher Education. Austin Doherty, then dean at Alverno College, and I led a session, "Assessing the Ineffables." In my judgment, moral and spiritual growth is not ineffable or "incapable of being expressed in words";[v] but it is an area in which assessments must especially be understood as "leaves in the wind," not as direct measures.

An Educational Resources Information Center [ERIC] search for research on "seminary students" and "values" brought up only the two following articles:

- "Human Nature and Authoritarianism in Seminary Students and Counselor Trainees." Mason, Robert; and others; Personnel Guidance J, 47, 7, 689-92, 69 Mar (EJ002625)
- "American Theological Students: A Survey of Their Value Commitments." Roscoe, John T.; Girling, Paul A., 1969 (ED027835)

A similar search for "medical students" and "values" brought up 79 articles. A review of the 79 indicated that not all appeared to deal directly with values and values' change, but the following titles taken from the 79 suggest that several do deal directly with values and values change:

- "Vanquishing Virtue: The Impact of Medical Education." (EJ631188)
- "Navigating the Wards: Teaching Medical Students to Use Their Moral Compasses." (EJ527868)
- "Toward a Person-Centered Medicine: Religious Studies in the Medical Curriculum." (EJ514952
- "The Effect of a Required Third-Year Family Medicine Clerkship on Medical Students' Attitudes: Value Indoctrination and Value Clarification." (EJ499647)
- "Sensitizing Residents to Moral Issues by Case Discussions." (EJ434655)
- "The Effect of Teaching Medical Ethics on Medical Students' Moral Reasoning." (EJ400941)
- "Value Orientations and the Effects of Professional Schools on Students." (ED214470)
- "Development and Validation of a Test Instrument for Assessing Value Preferences in Medical Ethics." (EJ259984)
- "Moral Reasoning as a Criterion for Admission to Medical School." (ED217817)
- "Interns' Attitudes and Values as Antecedents of Clinical Performance." (EJ198222)

It is interesting that these two ERIC searches found so many more research articles on medical students' values than on those of seminary students. (I ask any reader to let me know of a database that can give a fuller list of research articles on values change during seminary than I found in ERIC; my contact information is given at the end of the paper.)

Observation. Cultural anthropologists, by living among and observing a people group, can report the group's dominant values with considerable agreement among independent observers. I have wondered why we have not asked them to describe the pervasive values, attitudes, and work habits of student groups. In quality improvement efforts, organizations have identified "listening posts" or "observation posts" to listen or watch for repeated patterns of problems in processes. At Samford, the director of institutional research and I have begun an experiment with listening posts for student values, worries, and problems in living. We have identified certain individuals at strategic places to observe and reflect on student behaviors, such as the director of the health center, the university counselor, the associate athletic director, director of student ministries, etc. They will not report on any individual's behavior or identify anyone. Instead, they and we will work together to develop common definitions or categories of what they observe and report; and we will concentrate on recurring behaviors or patterns, not isolated, unusual behaviors.

To estimate beliefs, values, and attitudes from observed patterns of behaviors, observers must have opportunities to observe groups and individuals beyond formal events such as classes, chapel services, and called meetings. They must be able to observe in the candy store. That is, what do individuals do when they are free to choose? For example, if they could observe the materials students or former students select in airport bookstores and newsstands, English literature teachers could get some idea if they had affected students' literary tastes.

Obviously, one does not want to and will not spy on students in their private moments, but as Yogi Berra said, "You can observe a lot just by watching." This is particularly true if faculty and students live in close community.

The power of community became very clear to me in my five days at the Mundelein Seminary. Because Mundelein is a Roman Catholic seminary, the students all live in community with each other and their faculty, spiritual leaders, and many of the staff. Every student is a member of a 16-man *camerata* (known on campus as "cams") that lives in the same area of a resident hall with the cam leader, a priest with a suite in the same area. Students in each cam worship together, do sports and recreation together, and meet frequently with their cam leader as a group and individually. Every seminarian has a contact separate from his cam leader and spiritual director. While the spiritual directors function entirely within the "internal forum"-that is, the student's private, spiritual life is kept in absolute confidence-the contact functions in the "external forum." In the external forum, the contact observes the assigned seminarians, meets with each seminarian at least twice each semester, and collects information on the seminarian's progress from his faculty, his cam leader, and his parish-field work supervisor, but *not* his spiritual director. Every seminarian has a contact, and each contact is expected to report seminarians' progress every semester to the rector.

Living in community 24-7 allows a contact to observe patterns of behavior far beyond academic engagement in the classroom. Obviously, Protestant/Evangelical faculties do have similar opportunities to observe patterns of typical behavior on which they might base estimates of beliefs, values, and attitudes. At the same time, such faculties can learn a great deal by observing how students relate to one another, participate in worship and service activities, what they talk about over coffee or at lunch, and particularly when they are with them in off-campus retreats. If a faculty decided to learn what they could about students by unobtrusive observation and listening, they could become more accurate reporters and analysts by the following activities:

- Have someone skilled in ethnography train them in observation techniques.
- Work together as they observe and report to develop operational definitions and categories of behaviors to increase the chances they are observing and reporting the same things.
- At the end of each semester, faculty who often counsel students could report on recurring themes that they have heard that semester.
- A seminary faculty or administration might identify individuals on and off campus who interact with students outside class. Whether the British college or the American campus, if one wanted to know what is going on, one asked the janitors or custodians. Secretaries and other staff often hear and see what is all but invisible to faculty and administrators.

The greatest barrier to a faculty learning about students' values and how they may change during their seminary careers is not lack of knowledge or skill about systematic observation and unobtrusive indicators. It is deciding to observe and report, taking the time it requires, and learning from experience.

<u>Student Reports</u>. It often has been said, "If you want to know what people think, just ask them." If a faculty really wants to learn about students' spiritual development, why not ask them at a certain point each year to describe their spiritual struggles and practices? The faculty could craft the rubrics for the reports consistent with its shared theology and approaches to spirituality. In general, students will be more candid if they are not asked to identify themselves.

A strategy used at Samford might clarify what I am suggesting. All Samford freshmen take two semesters of Cultural Perspectives, an interdisciplinary course that integrates western and world civilization, history, and literature. To evaluate the effects of this two-semester course, on their first day of Cultural Perspectives in the fall, all freshmen were asked to write a brief essay on Plato's "Allegory of the Cave." At the end of the following spring semester, they were given a similar assignment. A Samford English professor, a Samford History professor, and a philosopher from another university read and compared the fall and spring essays of 80 randomly selected students.

One consistent difference emerged: On the first essay in the fall, students tended bring in their own beliefs and perspectives without entering into Plato's thought. In their spring essay, they wrote taking Plato's thoughts into consideration.

If seminarians wrote reports on their spiritual development each year using common rubrics, their papers could be compared from one year to the next. With a little help for individuals skilled in content-qualitative analysis, a faculty using at least two readers could identify patterns of change.

Values Inventories. Opinion and values inventories are used throughout higher education to evaluate programs and to identify prevailing student attitudes, to assess engagement in learning, to evaluate instruction and degree programs, to gauge satisfaction with administrative services and processes, etc. I was struck by the number of religious and values inventories (124) identified by Hill and Hood in *Measures of Religiosity*.[vi] Ralph Hood describes Gordon W. Allport's seminal work on intrinsic-extrinsic religiosity as the "dominating conceptual paradigm in the empirical psychology of religion during the last three decades."[vii] Therefore, seminaries using inventories to assess students' change in spirituality and values are in good company with their academic colleagues. Given the instruments described in *Measures of Religiosity*, seminaries appear to have a rich array of instruments available to assess religious values and spirituality.

Despite their convenience and frequent use, surveys by paper-and-pencil questionnaires present several problems:

- Low Response Rates. As most know, mailed questionnaires almost always have low rates of return, which raises certain questions. For example, did those who took the time to respond do so because they were unhappy and wanted to send a message? Or were they particularly pleased? Many techniques have been employed to boost response rates, such as keeping the questionnaire brief and simple, providing stamped/addressed envelopes, printing the questionnaire on brightly colored and odd-sized paper so that it is hard to hide on the desk, including token appreciations, and follow-up postcards and phone calls to nonrespondents. These efforts help somewhat, but response rates remain low. If a questionnaire is for students on campus, have students complete and return it during a required class or some other event where most, if not all, students attend. Return rates are about the same for Web-based surveys. Consider telephone surveys if the survey results are really important, but still only expect to get complete responses from one half of the sample.

- Positive Bias. Whether one is dealing with evaluations of individual performance, program effectiveness, or product satisfaction, people tend toward positive ratings. Scales such as the Likert 1-to-5 scale is especially vulnerable to positive bias. The 5 rating usually means "completely satisfied" or "outstanding" or "excellent," while the 4 rating usually suggests "some but not complete satisfaction" or "good but not outstanding." The 3 rating often represents "average" or just "OK." Given the pervasive presence of positive bias, some suggest discounting the 4 rating by 20%. A 3 rating on performance can often be interpreted as "unsatisfactory." When possible, it is generally better to use forced-choice or budgeted scales. A forced-choice scale might ask the respondent to rank eight to 10 characteristics of a program. For example, seminary graduates might be asked to rate their seminary experience by ranking the following outcomes from 1 (the most realized outcome) to 4 (the least realized):

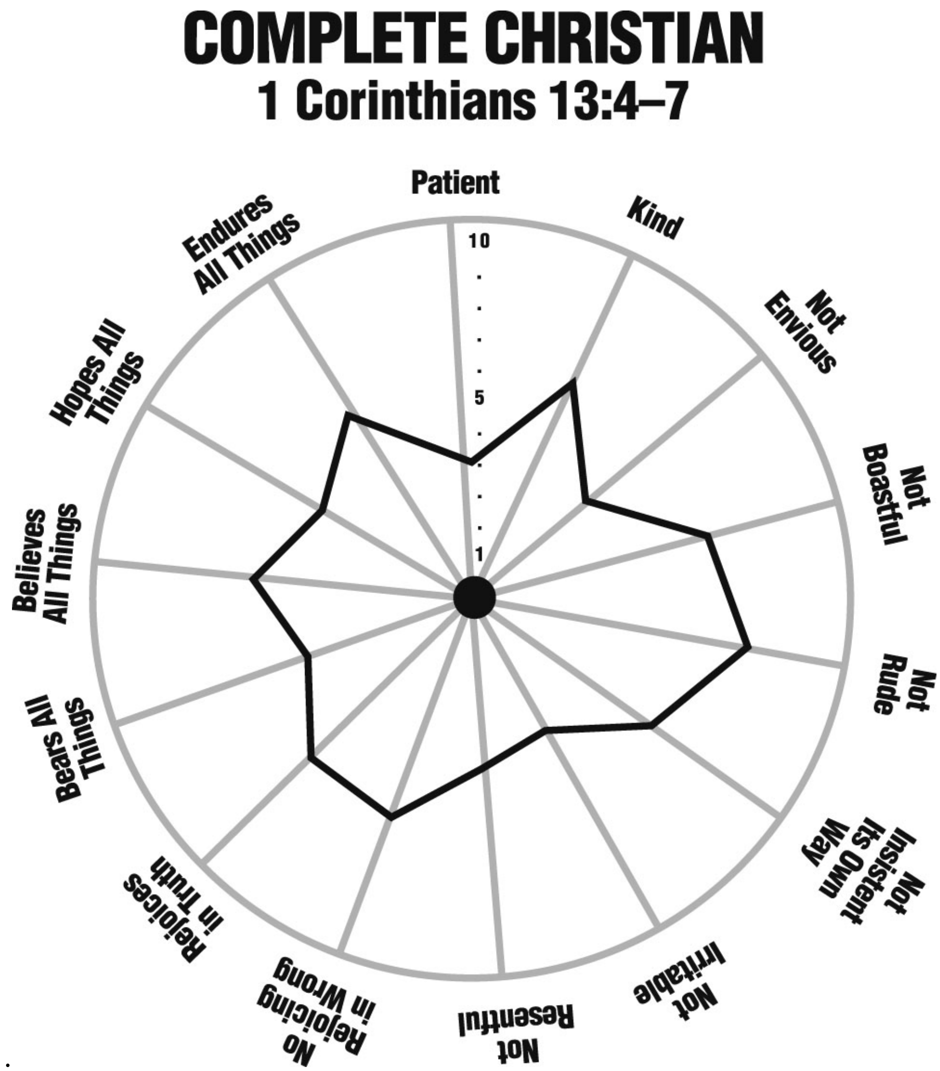
___ Knowledge of religious heritage
___ Understanding of cultural heritage
___ Growth in spiritual depth and moral integrity
___ Capacity for ministerial and public leadership

Or they might be asked to distribute 100 points to the four, giving the greatest number of points to the most realized or valued outcomes.

As suggested earlier, survey results should be used to stimulate reflection and focus tacit knowledge. They should not be taken as absolute or dismissed as ephemeral.

Quality Check. One of the quality improvement tools is the Spider Diagram. It purpose is to compare graphically ideal or perfect quality of a product or process to its current, actual quality. The graphic is composed of a circle with the center representing zero and the outer edge or perimeter representing 10. Each characteristic of the process or product is a spoke going from the center to the perimeter of the circle. Each characteristic of the process or product is rated one spoke from 0 (No quality.) to 10 (Ideal or Perfect).

It occurred to me that one could rate their Christ likeness. For example, the characteristics of Christ like love described in 1 Corinthians 13: 4-7 might be used as rating spokes. The following diagram is my effort to use the qualities of love described in this passage to represent the Complete Christian.




The inner "spider web" indicate estimates of how loving one is as compared to Paul's ideals. Seminarians might rate themselves each. If they wished, they could ask a mentor, spiritual director, or others who know them well to rate them. They could then compare how they rate themselves to how others rate them. If done each year, they could track their growth or decline.

Capacity for Ministerial and Public Leadership

I teach leadership courses in Beeson Divinity School's M.Div. and D.Min. programs, so I am tempted to wander into philosophies or theologies of leadership, but I spare the reader. Suffice it to say, if leadership is to be assessed, it must be defined in operational, observable terms. From my experience in and outside of seminaries, becoming specific about the kind of leader an institution hopes to produce will test most faculties' collective patience.

There are two obvious issues in assessing leadership. First, do students possess a working knowledge of a reasonable range of leadership styles, and are they able to compare and contrast them in terms of their effectiveness and congruence with Christian principles? Developing valid and accurate assessments of such understanding in and of itself would be a significant accomplishment. One possibility: Give students summaries of several well-known leadership styles, e.g. Homeric heroes such as Agamemnon and Achilles, Machiavelli, Jack Welch, and Lee Iacocca, and ask them to compare and contrast these leadership styles with that of Jesus.

To assess leadership behavior of any style requires something beyond academic tests. If one wants some idea of how a person is likely to respond in leadership, one place to begin is to examine past leadership behaviors. In general, past behavior is the best predictor of future behavior. The bio data prediction strategy described earlier would be one way to get this.

Another widely practiced approach is to put individuals in situations they are likely to face and observe whether they take a leadership role; if they do, observe how they attempt to lead. Situational testing has a long history, even into classical and Biblical times. Its more recent impetus came from spy-master Wild Bill Donovan when he led the cloak-and-dagger spy organization, Office of Strategic Services [OSS], in World War II, which later became the CIA. Spy candidates were put in situations such as those they would face in the field so that their master could to observe how they would respond. ATT adopted and adapted situational testing to corporate leadership using clinical assessment. Today, trained actors test medical students' diagnostic skills. The point in all this is that assessments of probable behaviors must approximate as closely as possible actual situations the individual may encounter.

While situational tests seem to be the best way to assess likely leadership, they are expensive to develop and administer. Here are some possible simulations for testing ministerial leadership that would have some of the value of situational testing but are less expensive.

<u>Cases</u>. Have trained case-study writers develop leadership cases with working ministers based on actual experiences. Then have the students respond to the cases.

<u>Video Vignettes</u>. Show students a video of a leadership dilemma often encountered by pastors and ask them to describe how they would deal with the situation.

<u>Public Arena</u>. Assign students to attend a local government meeting in which some policy is being discussed that affects the poor or otherwise disenfranchised. Ask them to describe the dynamics, and underlying values and agendas at work in the meeting, and how they might have contributed to discussion if their ministry was in that area.

## III. Assessment Feedback from Graduates and Others

Graduate Survey. Survey each year's M.Div. graduates about the effectiveness of the curriculum, pedagogy, and total experience in preparing them for ministry. A key question is, "What did you not get or not get enough of in seminary that you need in ministry?" Keep in mind the problems of survey return and positive bias that were discussed earlier.

Employer Survey. Survey employers of each year's M.Div. graduates about their ministry effectiveness.

Focus Groups. Focus groups comprised of graduates or employers can be useful, but it is important to use trained facilitators who are not connected with the seminary to get the most objective information.

Advisory Board. Appoint a Ministry Advisory Board of informed laypeople and pastors to review a summary of the results of the graduate and employer surveys to recommend curricular, pedagogical, and total experience changes.

The categories and items in ATS's Profiles of Ministry could be used to identify or develop the graduate and employer surveys. The $360^{\circ}$ assessment as employed in *Skillscope* and other managerial assessment instruments published by the Center for Creative Leadership might be adopted in the development of the employer survey.

## Concluding Reflections

Assessment is of little value unless it is part of a larger commitment to improve continuously. Within the context of continuous assessment's first task is to identify deficiencies; in academic assessment, its primary focus is on student learning. Are students learning what the educational program is designed to accomplish? Assessment's second task is to assist an inquiring faculty to find root causes of deficiencies. One of the most common mistakes is to attempt to improve on hunches without deep probing for root causes of learning deficiencies. Senge and associates advocate asking "Why?" five times before assuming a root cause has been identified.[viii]

Ted Ward observed that seminary faculties find it difficult to check empirically on praxis because they are so full of beliefs.[ix] He was not referring to their theological beliefs but to their educational beliefs. His observation resonated with me in my experience with faculty and academic administrators in general. It is interesting that those engaged in the most rational work find it particularly difficult to take a detached, objective view of their work. Admittedly, it is always difficult to reflect on the water in which one swims. That is why those interested in continuous improvement value the naïve observer's perspectives. A visitor is more likely to see the dust on the baseboards than the resident.

Assessment's only lasting value is to increase understanding of an organization or process for its improvement.

Russell Ackoff's following three sentences cut to the chase of this paper –

"An ounce of information is worth a pound of knowledge.
"An ounce of knowledge is worth a pound of information.
"An ounce of understanding is worth a pound of knowledge."[x]

John Harris
Orlean Bullard Beeson Professor of Education and
Associate Provost for Quality Assessment
Samford University (Retired)

---

[i] Harris, John and Sansom, Dennis. *Discerning Is More Than Counting (*Washington: American Academy for Liberal Education) http://www.aale.org/harris-sansom.pdf (18 December 2003).
[ii] Timothy George, dean, Beeson Divinity School, Samford University, December 19, 2003, personal communication.
[iii] Lee J. Cronbach, *Essentials of Psychological Testing* (New York: Harper & Row, Publishers, 2nd Ed., 1960), 29-34.
[iv] Jennifer Crafts, "Using Biodata as a Selection Instrument." ERIC Database # ED338702
[v] Webster's Ninth New Collegiate Dictionary, s.v. "ineffable."
[vi] Hill, PC and Hood, RW, *Measures of Religiosity* (Birmingham, Alabama: Religious Education Press, 1999).
[vii] Ralph W. Hood, Jr., "Intrinsic-Extrinsic Religiosity," William H. Swatos, Jr. Ed., *Encyclopedia of Religion and Society,* Hartford Institute for Religion Research, Hartford Seminary**,** Hartford, CT. <http://hirr.hartsem.edu/ency/allport.htm> (18 December 2003).
[viii] Peter M. Senge, *et al*. *The Fifth Discipline Field Book (*New York: Doubleday, 1994), 108-112.
[ix] Personal communication with Ted Ward, October 31, 2003.
[x] Russell L. Ackoff. *Ackoff's Fables: Irreverent Reflections on Business and Bureaucracy (*New York: John Wiley & Sons, Inc., 1991), 75.