Performance Theory Based Outcome Measurement in Engineering Education and Training

William E. Dillon, *Member, IEEE*, George V. Kondraske, *Senior Member, IEEE*, Louis J. Everett, *Member, IEEE*, and Richard A. Volz, *Fellow, IEEE*

***Abstract*—An approach is presented to improve engineering education that is based on new concepts of systems performance and classic feedback theory. An important aspect is the use of General Systems Performance Theory (GSPT) to provide a performance model of the educational system and as a basis for the key outcome metrics: the volumes of performance capacity envelopes of individual students. Feedback is aimed at achieving both better curriculum design and teaching methods. In addition to conceptual issues, a web-based implementation plan and experimental validation plan is described. The quantitative modeling approach taken, including choice of appropriate levels of abstraction, has provided better understanding of the system used to provide engineering education and a basis for quantitatively linking components of the program to student performance in a causal manner. The educational system performance model is discussed in the context of competency models. It is believed that this approach holds promise for not only documenting a meaningful type of outcome but also for providing insight into the rationale for steps taken in attempts to improve an educational system.**

***Index Terms*—Assessment, education, outcome, performance, performance capacity envelope, performance theory.**

I. Introduction

THE COMPLEXITY of engineering curriculum design (including tradeoffs between "fundamentals" and "applicable skills") is increasing and new avenues are emerging for education delivery such as distance learning, web-based and supplemented courses, etc. [1]. From a macroscopic viewpoint, educational programs run in largely an "open-loop" fashion. Feedback *from* students is minimal, generally "too late," and microscopic (e.g., "course" evaluations). Feedback *about* students (especially *how well* they integrate knowledge) other than from isolated courses is virtually nonexistent. The increasing cost, complexity, and importance of education to the nation has resulted in quality scrutiny. These collective factors have motivated an emphasis on "outcome" [2]–[4], in attempts to close gaps between actual and desired educational program performance. It is strongly felt that the combination of traditional and new systems engineering methods provide a compelling basis for realizing progress in improving educational effectiveness.

Manuscript received August 15, 1999; revised November 30, 1999. This work was supported in part by the Texas Higher Education Coordinating Board, Advanced Technology Program and United Space Alliance, Inc., Houston, TX.

W. E. Dillon and G. V. Kondraske are with the Human Performance Institute, University of Texas, Arlington, TX 76019-0180 USA (e-mail: dillon@uta.edu).

L. J. Everett is with the Mechanical Engineering Department, Texas A&M University, College Station, TX 77843-3128 USA.

R. A. Volz is with the Computer Science Department, Texas A&M University, College Station, TX 77843-3112 USA.

Publisher Item Identifier S 0018-9359(00)04313-2.

It is evident that there has been a growing general consensus that "outcome" is very important in education, as well as other areas such as health care. However, there is much less agreement about what outcome is and how it should be measured. In education, when one talks about outcome one is talking about students and more specifically, their ability to perform in a given field of expertise. Thus, to better understand "outcome" it is asserted that it is essential to better understand *performance*. The concept of performance pervades nearly all aspects of life, especially decision-making processes associated with human and man-made systems. Yet, it is not well understood theoretically and approaches to its modeling and measurement have been largely *ad hoc*, regardless of the selected field. Despite this, neither practitioners nor researchers in education can be overheard saying, "What we need is a good, comprehensive systems performance theory . . . ." It is argued, however, that this could be part of the answer to general problems surfacing as specific problems in different fields (such as education) that pertain to systems, tasks, their interface, and the concept of "performance." This paper addresses the goal of outcome assessment from the perspective of performance by using a relatively new body of work known as General Systems Performance Theory (GSPT). Here, the key systems of interest are an "education system" and "the student." At issue is insight into a cause-and-effect, quantitative understanding of the interface of these systems to the tasks they execute.

This work described is a component of a growing effort to provide a more robust and scientifically sound conceptual basis for improving the quality education programs that builds on the authors' collective experiences and findings from the past 12 years. Previous work in human performance, systems performance theory, and the application of these to training systems, coupled with the authors' experience as educators, have produced insights into generalizations regarding the contemporary educational process. This work is now being extended to suggest a more rigorous, quantitative, and performance-oriented basis for modeling "educational systems," defining and measuring "educational outcome," and improving educational program effectiveness.

II. Background: General Systems Performance Theory

Although a considerable body of material known as general systems theory exists, the concept of performance has not been incorporated in it nor has performance been addressed in a general sense elsewhere to the best knowledge of the authors. Most knowledge that does exist about performance and its quantitative treatment has evolved within specific applications, where generalizations can easily be elusive. Performance is multi-faceted, pertaining to how well a given system executes an intended function and the various factors that contribute to this.

GSPT developed by Kondraske provides a relatively new set of robust modeling constructs to explain how to model and measure all aspects of system performance, characterize tasks, and understand the interface of systems to tasks. GSPT was initially presented [5] in the context of developing a model (the Elemental Resource Model or ERM) for human performance and human-task interfaces. Following initial presentation, it emerged as an entity separate from the ERM. While some modest refinements in both GSPT and the ERM have been made [6]–[9], the basic approaches, terminology, and constructs used have remained quite stable. GSPT has been used to provide a conceptual basis for measurement and assessment in rehabilitation [10], to define work site modification for individuals with disabilities as an engineering process [11], to compute composite measures of performance in clinical drug trials [12], to objectively address gender issues in physical education and sport [13], to provide a quantitative measure of motion quality [14], and to model and measure telerobotic system performance [7]. Applications in other areas are in progress. It has also been incorporated into the work of others [15]–[17]. Experimental studies designed to evaluate key constructs of the ERM have been carried out in the areas of vocal performance [18], upper extremity motor control [19], and characterization of neuromuscular systems [20]. It has also been employed to derive a new method for task analysis and performance prediction called Nonlinear Causal Resource Analysis which has been demonstrated experimentally in the context of human locomotion [21] and assessment of professional athletes [22].

A key aspect of GSPT is the exclusive use of a "resource" construct to model all aspects of system performance. Systems are characterized as "possessing" different types and amounts of *performance resources* (e.g., speed, accuracy, user friendliness, etc.), each of which becomes a target for quantification efforts to determine "resource availabilities." Each performance resource represents one dimension of a multidimensional performance space (i.e., a dimension of performance, or DOP); a goal of system characterization is to determine the "performance capacity envelope" (PCE) for each subsystem. The powerful notion of a performance envelope first emerged in the aerospace industry. "Pushing the envelope" has become a widely employed, but loosely defined, notion. The dimensions of performance and metrics used with them in the aerospace world (i.e., speed, altitude, and range) naturally lead to a performance envelope. Such is not the case for many other systems. However, GSPT teaches how to achieve a performance envelope for any system.

The visualization of a system in terms of its performance envelope greatly facilitates consideration of the interface of that system to a task. Any task can be modeled in terms of the demands it makes on each of the systems performance resources. Thus, each task is represented by a point in a multidimensional performance space in which the performance capacity envelope is defined. If the point lies inside the envelope, that system *possesses* the performance resources necessary to accomplish that task. If the point representing a task lies outside the envelope that task cannot be accomplished by that system; i.e., one or more performance resources are *limiting resources*.[1]

Given this perspective, it can be readily seen that a performance capacity envelope with a larger volume contains more than one with a smaller volume. This can be interpreted as that system possessing a greater capacity to accomplish tasks; i.e., it's envelope encloses more tasks. In an n-dimensional performance space with each DOP properly defined using the resource construct, the overall performance capacity is simply measured as the volume enclosed by the envelope; i.e., the product of all the constituent performance capacities, assuming an idealized (rectangular versus curved) envelope. Another way of interpreting this is from the perspective of joint probability. Assume that the measure along each dimension is a properly normalized probability (p_i) representing the probability that the system in question possesses enough of the performance resource along DOP "i" to accomplish task "k." If the task makes demands on both performance resources simultaneously (which is typical), the probability of having sufficient performance resource availability along DOP 1 and DOP 2 is $p_1 \times p_2$ (assuming independence, which will not always be a perfect assumption). These notions readily extend to situations in which an "n-dimensional" performance space is considered.

III. Application to Education

A systems engineering approach is taken that is centered around the development of quantitative systems performance models and applied to two hierarchical levels:

1) the overall educational system of interest, and
2) major constituent subsystems.

Also, the context for discussion is engineering education, although concepts are generally adaptable to any educational or training program.

For the present purpose, the product (or output) of the educational system that is of interest is "the student." To facilitate the GSPT view, it is helpful to view humans (i.e., students) as adaptable, reconfigurable systems. Applying GSPT methods, one first identifies the function associated with this potentially multifunctional system that is to be considered. While it is possible to do this at different levels of abstraction, a view is selected from the broadest level for illustrative purposes and it is considered that the function of interest is "to do engineering" (i.e., engineering tasks). Thus, this functional consideration results in the student being viewed as operating in a mode in which they are "configured" to execute the function of an engineer. This process facilitates dealing with the vast complexity of human systems, allowing a focus on one function (i.e., one configuration) at a time. Separating this "student system" from the educational system, focus is initially restricted to:

1) a system with some capacity to execute engineering tasks (i.e., the student) and,
2) the engineering tasks themselves.

The notion of the performance capacity envelope, as presented above and illustrated in Fig. 1, is now relevant. It is asserted that

[1]In this sense, the mathematics of performance and system-task interfaces is similar to that of chemical reactions.

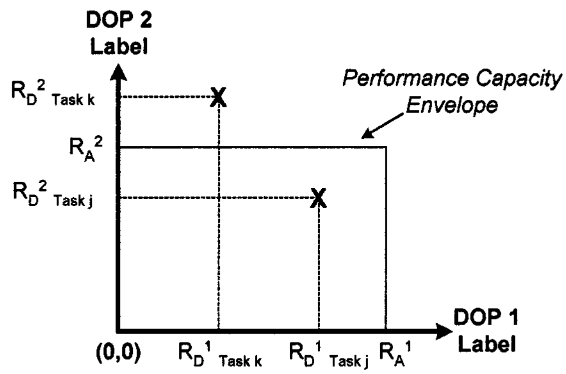

Fig. 1. GSPT-based characterization of 1) a system with two dimensions of performance (DOP's) considered and an idealized performance capacity envelope (representing performance resource availability) and 2) two different tasks as points "X" (representing performance resource demands). Assessment of the interface of this system to these tasks (i.e., Can this system successfully accomplish Task A?) involves determining whether the point lies inside or outside of the envelope. Note the importance of the resource construct in defining dimensions of performance. An envelope is not obtained unless 1) the dimension labels represent desirable quantities (e.g., speed versus response time) and 2) the process used to measure each quantity produces a larger numerical value when "more" of that resource is available.

a student's engineering performance capacity envelope is "the" major determinant of overall educational system effectiveness and therefore the key outcome metric of interest.

The next objective is to define the performance capacity space in which a given student's PCE can be estimated. This process is tied to the choice of the level of abstraction selected in identifying the function of interest to be considered. For example, it is possible to consider a more narrow function (e.g., "design of low frequency passive filters") than that selected previously (e.g., to "do engineering"). This choice is obviously important in determining the dimensions of performance that form the performance space in which the performance capacity envelope will be determined.[2] Thus, for the purpose of proceeding with the presentation of key conceptual aspects of the proposed method (while maintaining a healthy respect for issues raised), a simplified model of the student as an engineer will be assumed (using a method commonly used in teaching).

The process of defining the performance space of interest amounts to defining the constituent dimensions of performance (DOP's).[3] GSPT teaches that DOP's represent "unique qualities of a system that contribute to how well that system executes its function." In the present context, one can begin by selecting relatively independent skill areas and associating each with a DOP. For example, consider electrical engineering (EE) and mechanical engineering (ME) students. A set of DOP's that might be employed for outcome determination at various stages of the respective student's progress are:

1) digital logic and electronic design (for EE);
2) solid mechanics and vibration (for ME);
3) circuit theory, control system, dynamics, and thermodynamics (for EE and ME).

Fig. 2 brings together a number of these concepts in the context of engineering education. Illustrative engineering PCE's for three different students are shown, postured against the same engineering task (i.e., one that makes demands for certain "amounts" of circuit theory, control systems, and electronic design skills). Re-emphasizing an important GSPT concept, the PCE volume represents an important single number outcome metric that is interpreted as the capacity of the student to accomplish tasks that make demands on the skills that form the performance space. This is so because tasks are viewed quantitatively in terms of "how much" of each skill (i.e., performance resource) is demanded in order to achieve a given level of performance in that task. These numbers define a point in the relevant performance space. A larger student PCE contains more "points" (tasks) and therefore the student has a greater capacity to execute tasks of this type. It is clear also that the same volume can be achieved in an infinite number of ways. Thus, when the task is not clear (i.e., a more "general purpose" view is desired or appropriate), the volumetric performance capacity is appropriate. If a specific task or set of tasks with known demands is being considered, the performance capacity along each dimension must be individually considered. However, it is noted that if the product of task demands exceeds the PCE volume, it is certain that the student will not be able to accomplish the task. The multidimensional PCE volume can therefore be a useful screening measure.

Next, an education program (e.g., such as one designed to produce BSEE's) is considered as a system that (ideally) produces people with "appropriate" PCE volumes. Viewed from a particular level and type of abstraction, important subsystems within this system are the curriculum (composed of lower level systems or courses) and the faculty (composed of lower level systems called instructors). While space prohibits detailed treatment, attention is now directed to the PCE's of these subsystems and the overall educational system. It is argued that, in a causal view of the educational process, it makes intuitive sense that the ability to achieve the goal of producing students with appropriate PCE's is dependent on the respective PCE's of courses and instructors. It may be less intuitive, however, to base the PCE's of courses, instructors, and the overall educational program on student performance capacities. But this latter thought is central to the motivation behind the proposed outcome-based program assessment; i.e., the educational program performance (and its components) are assessed by observing the outcome as defined by student PCE's. Does it make sense to measure the capacity of a circuit theory course to impart circuit theory skill in students by measuring the circuit theory performance capacity of students who have been through that course? Different students follow different paths in a typical educational program and may be exposed to different instructors, as well as different textbooks with different topics (although the course number may

[2]There is much work to be done here in applying these concepts to any given curriculum and detailed discussion of approaches and alternatives is beyond the present scope. Nonetheless, while this process is recognized as challenging and will likely have controversy associated with choices made, it is believed that it can be accomplished with reasonable fidelity and acceptance through use of well-known techniques such as the Delphi method, It is also recognized that this involves much of the same decision-making that is currently a necessary part of curriculum design (i.e., what should students in a given curriculum or given course know and know how to do?).

[3]The selection of DOP's here is not intended to comprehensively cover the full curriculum of a program, but rather representative segments thereof. The intention here is simply to illustrate the steps and general process.

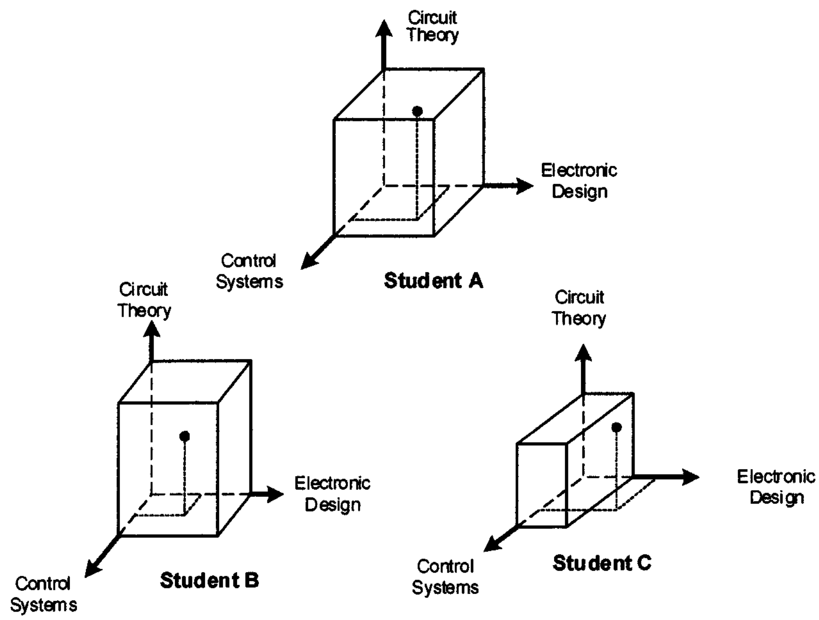

Fig. 2. Outcome expressed as engineering performance capacity using PCE's for three different students. The same engineering task (i.e., a specified point in performance space requiring "x" amount of electronic design skill, "y" amount of control systems skill, and "z" amount of circuit theory skill) is also shown in each plot. Students B and C can be seen to have performance resource deficiencies (of different types); they will not be able to accomplish the selected task with their present performance capacities—but for different reasons. Note that the point representing the task lies outside of their respective PCE's. In contrast, the point representing the task lies within Student A's PCE.

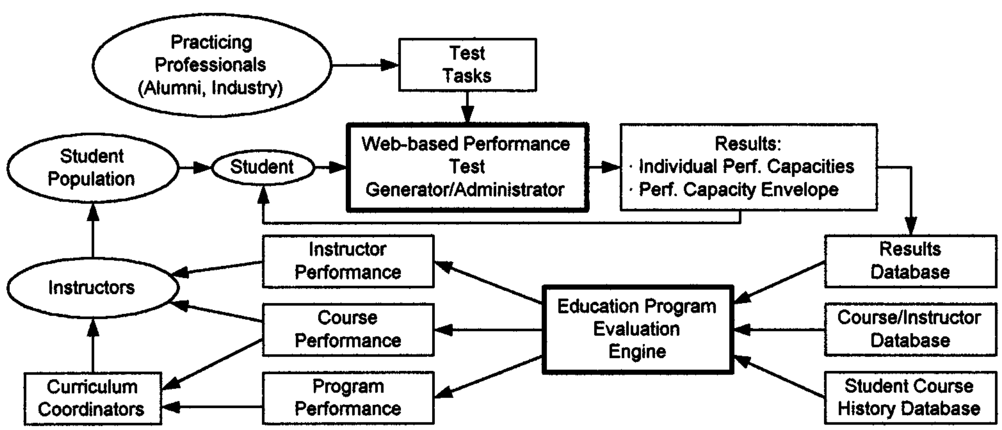

Fig. 3. Systems performance model of an educational program, including components that integrate performance-based outcome measurements into a closed loop feedback scheme to achieve improvement in overall program effectiveness.

be the same). In addition, more than one course may contribute to a given skill area as a student progresses through the curriculum. Thus, it is argued that the notion of determining instructor, course, and program performance from student PCE's makes good sense and that measurements should be made on a program level (e.g., end of sophomore year, etc.) as well as on a course level (i.e., instructors can adopt the PCE approach in course design and test formulation).

In Fig. 3, the systems performance modeling concepts introduced above are applied to an overall educational system and its subsystems. Also, subsystems associated with PCE measurement have been integrated in a manner that results in a classic closed loop feedback control model in which student performance capacity (i.e., the average of all students PCE's, perhaps estimated by a representative sample of students) is the controlled variable. The key concepts of the process represented are as follows:

1) program outcome is measured in terms of student PCE's determined through a testing process that is independent of individual courses;
2) individual student PCE's (composite and for each constituent DOP) are mapped to courses and to instructors in order to create data sets for each of these entities (i.e., using course instructor, student history, test result databases, and "rules"[4] that define the contribution of a given course to different envelope DOP's);
3) data sets (individual student PCE's) are weighted when appropriate (e.g., based on the amount of contact a given

[4]These rules would likely be based on analysis of course syllabi to estimate the relative contribution of a given course to a given DOP.

student had with a given instructor) and processed statistically to estimate course, instructor, and program performance;

4) this information is made available to the appropriate education professionals (i.e., feedback is achieved);
5) where performance does not meet established criteria, changes are made that would be expected to lead to improvement.

IV. Implementation Plan

It is planned to implement and test the proposed approach in an incremental fashion. The process will start with a reasonable representation of the overall process represented in Fig. 3 within one department (e.g., electrical engineering) and include only a subset of all skill areas. It is hoped that this will be used to educate faculty, administrators, and students to enlist support for expansion to include other skill areas and other departments.

A. Web-Based Performance Test Generator/Administrator

Referring to Fig. 3, it is proposed that a Web-based Performance Test Generator/Administrator will be used to determine student PCE's. The test generator will produce an "engineering performance capacity" evaluation for a student who logs on to a specified web site. A "test" for a given dimension of performance (e.g., circuit theory) will consist of a set of challenges randomly selected from a carefully generated library of "test tasks" designed to stress students along the respective dimension. Ideally, these test tasks should be formulated by professionals who are *external to the education program*. It is anticipated that alumni and industry representatives of departmental advisory boards will be willing, if not eager, to participate in this process. The "test tasks" will also be designed to facilitate automated performance scoring. Each participating student will be administered tests for each of the dimensions listed in Stage 1. The collective results will determine the student's performance capacity envelope. As per GSPT, a single number final score will be computed as the mathematical product of the capacities demonstrated along each dimension of the relevant performance space.

B. Educational Program Evaluation Engine

A special software package is envisioned that will serve as a basic "evaluation engine." Briefly (referring to Fig. 3), this will utilize databases (student performance capacity, course/instructor, and student course history) to determine—for a given student performance capacity associated with a specific dimension of performance—a numerical score contributing to the course and instructor performance metrics. This score will in fact be the respective dimensional student performance capacity (i.e., the "outcome"). Note that one-to-one mappings between dimensions of performance, courses, and instructors are *not assumed*. For each measurement period, actual course and instructor performance metrics will be determined as statistics (mean, median, and standard deviation) of the appropriately mapped student performance capacities.

Overall program performance (or a "limited" version thereof, due to initial scope limitations) will be computed as the *mathematical product* of the averages of student performance capacities for each dimension of performance (e.g., circuit theory) in the respective program (e.g., UTA-EE, A&M-EE, etc.). Note that a low score along any one dimension will result in a low overall program score. Thus, this method will readily bring to light major problems in any one area. It is also planned to explore the use of weighting factors (determined by accepted methods of obtaining expert consensus) in determining program, course, and instructor performance metrics. However, the main inputs are always student performance capacities.

C. Experimental Plan

Experiments are proposed that test the basic hypothesis that educational program quality can be improved through use of the proposed system. It is unrealistic to expect that this hypothesis can be definitively answered in a short period of time with data from only one implementation site. There is confidence, however, that sufficient data can be gathered, to *support* acceptance or rejection of the hypothesis. For example, it is proposed that the program be started by recruiting a minimum of 50 junior and senior UG engineering students at Texas A&M and UT Arlington (with an approximately equal EE/ME mix). They will take a web-based performance capacity test during each of at least three semesters ($>$total of 150 "outcome" observations). The DOP's to be included would be similar to the representative set described above. At present, the system delay time is one semester.

The student PCE's will be processed by the initial version of the evaluation engine (which is expected to evolve in sophistication) as described above and utilized as feedback as shown in Fig. 2 to a subset of "involved" course coordinators and instructors (the "treatment group"). The remaining subset will comprise the "control group," who will not receive feedback. Thus, in addition to individual student performance capacities (overall and for each DOP), data will be available that represents course and instructor performance in "treatment" and "control" groups for each of three semesters. One expects to find that, where "room for improvement" exists, treatment group course and instructor performance metrics (composites of student performance capacities) for second and third data collection semesters will exhibit greater gains relative to the control group.

V. Discussion

In this paper, the use of GSPT concepts to develop an outcome-oriented system performance model for an undergraduate engineering education system is described. Outcome-oriented methods have traditionally focused on items that are the products of an individual's global "performance" in a particular endeavor [23]. If applied in the traditional manner to engineering education, this would require observations of former students in an actual work environment. Feedback obtained from such observation may indeed relate to educational program effectiveness, but it is likely that the educational program in place at the time such feedback was available is not the same that produced the observed individuals. Generally speaking, classical outcome metrics are not directly connected to particular skills (e.g. salary, number of projects completed, etc.) and may be based on qualitative assessments. In fact, the qualitative character of most outcome-oriented methods is often viewed to be a strength of the method [24]. The authors' view of outcome attempts to fill a perceived void between such traditional outcome metrics and the data that is normally gathered from students while they are in residence in an academic program. One finds this view of outcome to be in good agreement with that reflected in new ABET standards [3]. In this sense, one focuses on the "more immediate outcomes" at different segments of the educational process (e.g., at or near the end of sophomore and senior years) where not only the knowledge and skills from different courses, but also their integration by the student, are of interest.

Some elements of the proposed approach are perhaps more similar to competency-based education (CBE) methods [25]–[29]. The essence of this approach is attributed to McClelland [30] who advocated determining what leads to superior performance by identifying top performers and finding out what they do. Since introduced, competency-based methods have gained widespread use in human resource management as well as in elementary and secondary education, but less so in higher education. However, implementations under that label do not always faithfully replicate the original notion and there exists what may be termed as different variations of CBE. The cited works on this topic contain a good deal of methodology (e.g. for identifying competencies in a given area) that can be applied to facilitate implementation of the proposed performance theory based approach. While it is has been argued by some advocating true outcome-based methods that the mere presence of competencies does not guarantee that they will be used, it is felt that there is considerable merit to establishing the level of certain competencies at selected points in an education program. Clearly, if certain levels of skill are not present, there will be no possibility that the desired workplace behaviors can be observed. In this sense, the student PCE's are viewed as establishing necessary but not sufficient conditions for success.

In comparison to CBE methods, the proposed approach focuses on identifying skill areas, or dimensions of performance, (which can be viewed as being analogous to competencies) and determining how well a representative sample of individuals execute in that area (i.e. versus what top performers or any other individuals do). Like the system advocated here, most CBE methods involve task analysis as the means used to identify the competencies (e.g., spelling, grammar, telephone skills for a secretary). However, these competencies are generally considered only in binary form (i.e., "required" or "not required," from the task perspective; "present" or "absent" from the human perspective) or in a normative referenced way (e.g., below standard, standard, above standard). In the GSPT view, each skill area is a continuous dimension to which a selected level of measurement resolution can be applied. Furthermore, the CBE method includes no provision for combining metrics across different competencies to obtain a single number "overall performance" for the individual such as the volume of the PCE as has been described. Returning to the control system aspect of the model, GSPT provides the basis for the computation of a meaningful quantitative variable to be controlled (i.e., the average of the volumes of a representative sample of student PCE's).

Given this, there are also more substantial fundamental differences compared to other methods proposed or in use. First, it is noted that GSPT is applied to obtain a complete systems performance model of an "educational system" as the primary system of interest and not only to the individual being trained. The student is modeled as a subsystem within this system. These two systems execute different functions (i.e., production of engineers and the actual "doing" of engineering) and one addresses how well each function is executed in terms of PCE's and links them quantitatively. Furthermore, the use of GSPT results in a framework to quantitatively link job tasks (measured in terms of the amount of given performance resources required) to the qualities of the human that executes them (measured in terms of the amount of given performance capacities available). This causal type, threshold dependent linkage is not only intuitively appealing because resource economics is pervasive in everyday life, but has also been experimentally validated in a number of performance-related contexts as noted earlier.

Further insight into fundamental differences offered the GSPT-based approach is evidenced by a definition of "a competency" produced from the collective inputs of several hundred experts at a 1995 competency conference [27]:

> "a cluster of related knowledge, skills, and attitudes that affects a major part of one's job (a role or responsibility), that correlate with performance on the job, that can be measured against well-accepted standards, and that can be improved via training and development."

Here, attention is called to the expectation present in the general community of training experts that the level of competency available will correlate with level of performance on the job. It has been argued rather extensively elsewhere [13] and demonstrated through careful review of experimental data using GPST constructs as a guide that it is a conceptual flaw to expect such correlations. This is due to the threshold nature of performance resource economics and the notion that the level of success achieved in a higher level tasks (e.g., job performance) depends on the logical combination (i.e., AND function) of lower level performance resources. Multivariate regression methods that have been the traditional tools of predictive research ignore the presence of thresholds (i.e., nonlinearities). Perhaps the best evidence speaking against the widely held correlation belief is that there are no quantitative correlation-based models in widespread use for predicting level of job success. Other works within relevant fields demonstrate a long-standing need for a true rigorous framework for dealing with performance [27], [31], [32]. This is perhaps most simply evidenced by the absence of terms such as "capacity," "envelope," and "dimension" in the index section of each of the cited texts and the absence of any quantitative expression that attempts to mathematically related system characteristics to task requirements. In Landy *et al.*, a history of performance measurement (primarily related to humans) is provided that illustrates the hunt for firm ground. Lastly, it is considered noteworthy that there appears to be no single or several seminal works that are consistently cited in training and education literature that deal with the matters addressed here with GSPT and a performance modeling approach. It is suggested that the GSPT-based educational system performance model is novel, can augment and provide further insight into methods such as CBE, and also help address other enigmatic and long-standing issues in training and education.

By integrating relatively new theory with current needs in a real-world context, the engineering of solutions to anticipated implementation issues will no doubt be forced. Many of the steps required for implementation are only briefly addressed in this presentation. While some are straightforward, others such as defining DOP's for a given program and development of tests and metrics that validly reflect the amount of a given performance capacity present will no doubt be more challenging. Reliability and validity of web-based PCE determinations must also be investigated and established. In understanding the educational system, one should consider "taking a page from our own book." In teaching circuit theory, one does not begin with a diagram for a television or radar system, but rather a circuit with one source and one resistor. In teaching statics, the first problem considered is often a simple cantilever beam and not a complex suspension bridge. These "simple cases" are essential to make clear the principles of nature at work and provide the basis for systematic, structured modeling and evaluation of more complex systems in which the same principles apply. A similar, incremental approach is suggested for educational system performance model development. With some obvious cautions, an analogy can be made in which an educational system is viewed to have aspects that are similar to a manufacturing system. With a performance model such as that proposed as a starting point, there are substantial, proven analytic tools from the manufacturing world [33]–[36] that can be readily adopted to help with the process of achieving improved educational quality (i.e., "product" quality). In any case, full implementation of the proposed methods cannot occur as the result of any single project in an "overnight" fashion. Rather, it is viewed as adding a missing component to an educational program that requires a shift in the way of doing business. One would anticipate first a recognition and understanding of the general framework and principles, followed by a gradual but steady incremental development and incorporation of various elements.

One notes that the approach taken is particularly well suited to using the same outcome measurement tools (i.e., web-based PCE tests) across different universities. In addition to improved student performance capacities, another benefit is the potential for lowering educational costs as related to efforts aimed at optimizing program content [37]. If sufficient redundancy can be trimmed from engineering curricula to allow average students to graduate in four years rather than 4.5–5 years, the cost of producing engineers can be reduced as much as 20%. It is also anticipated that interest from industry in using assessments of student performance-based outcome will help guide the curriculum development process and also justify decisions made. The needs addressed are critical in, but not unique to, engineering education. However, they are also present in educational programs associated with any discipline. The performance theory-based approach described is considered to be general and therefore applicable to any educational program. It is hoped that these views generate thought, critique, and perhaps action. Interest is welcomed from those who would like to collaborate in any of the many aspects required for prototype implementation, development of various components, and experimental studies.

REFERENCES

[1] "Special issue: Applications of information technologies to engineering and science education," *IEEE Trans. Educ.*, vol. 39, Aug 1996.

[2] *Criteria for Accreditation*, 11th ed: Southern Association of Colleges and Schools, 1997.

[3] *Engineering Criteria 2000*: Accreditation Board for Engineering and Technology, 1997.

[4] J. A. Shaeiwitz, "Outcomes assessment in engineering education," *J. Eng. Educ.*, pp. 239–246, July 1995.

[5] G. V. Kondraske, "Human performance: Science or art?," in *Proc. 13th Northeast Bioengineering Conf.*, K. Foster, Ed., 1987, pp. 44–47.

[6] ——, "Quantitative measurement and assessment of performance," in *Rehabilitation Engineering*, R. V. Smith and J. H. Leslie, Eds. Boca Raton, FL: CRC, 1990, pp. 101–125.

[7] G. V. Kondraske and G. J. Khoury, "Telerobotic system performance measurement: Motivation and methods," in *Cooperative Intelligent Robotics in Space III*. Bellingham, WA: SPIE (Int. Society for Optical Engineering), 1992, vol. 1829, pp. 161–172.

[8] G. V. Kondraske, "An elemental resource model for the human-task interface," *Int. J. Technology Assessment in Health Care*, vol. 11, no. 2, pp. 153–173, 1995.

[9] ——, "A working model for human system-task interfaces," in *The Biomedical Engineering Handbook*, J. Bronzino, Ed. Boca Raton, FL: CRC, 1995, pp. 2147–2164.

[10] ——, "Measurement tools and processes in rehabilitation engineering," in *The Biomedical Engineering Handbook*, J. Bronzino, Ed. Boca Raton, FL: CRC, 1995, pp. 2108–2124.

[11] ——, "Workplace design: An elemental resource approach to task analysis and human performance measurements," in *Proc. Int. Conf. Association for the Advancement of Rehabilitation Technology*, 1988, pp. 608–611.

[12] K. Syndulko, W. W. Tourtellotte, R. W. Baumhefner, G. W. Ellison, L. W. Myers, G. Belendiuk, and G. V. Kondraske, "Neuroperformance evaluation of multiple sclerosis disease progression in a clinical trial: Implications for neurological outcomes," *J. Neuro. Rehab.*, vol. 7, no. 3, pp. 153–174, 1993.

[13] G. V. Kondraske and P. J. Beehler, "Applying general systems performance theory and the elemental resource model to gender-related issues in physical education and sport," *Women Physical Educ. Sport J.*, vol. 3, no. 2, pp. 1–19, 1994.

[14] C. A. Fischer and G. V. Kondraske, "A new approach to human motion quality measurement," in *Proc. 19th Annu. Eng. Medicine Biol. Soc. Conf.*, 1997, pp. 1701–1704.

[15] H. P. Frisch, "Man/machine interaction dynamics and performance analysis," in *Proc. NATO-Army-NASA Advanced Study Institute on Concurrent Engineering Tools and Technologies for Mechanical System Design*. New York, 1993.

[16] T. G. Mayer and R. J. Gatchel, *Functional Restoration for Spinal Disorders: The Sports Medicine Approach*. Philadelphia, PA: Lea & Febiger, 1988, pp. 66–77.

[17] R. A. Cooper, *Rehabilitation Engineering Applied to Mobility and Manipulation*. London, U.K.: IOP, 1995.

[18] M. Jafari, K. H. Wong, K. Behbehani, and G. V. Kondraske, "Performance characterization of the human pitch control system: An acoustic approach," *J. Acoust. Soc. Amer.*, vol. 85, no. 3, pp. 1322–1328, 1989.

[19] G. V. Kondraske, "Neuromuscular performance: Resource economics and product-based composite indices," in *Proc. 11th Annu. IEEE Eng. Medicine Biol. Soc. Conf.*, 1989, pp. 1045–1046.

[20] P. J. Vasta and G. V. Kondraske, "An approach to estimating performance capacity envelopes: Knee extensor system example," in *Proc. 19th Annu. Eng. Medicine Biol. Soc. Conf.*, 1997, pp. 1713–1716.

[21] ——, "Performance prediction of an upper extremity reciprocal task using nonlinear causal resource analysis," in *Proc. 16th Annu. Eng. Medical Biol. Soc. Conf.*, 1994, pp. 305–306.

[22] P. Beehler, P. J. Vasta, G. V. Kondraske, and B. Ward, "Motor performance assessment using a resource economic strategy: Concepts and demonstration with NFL quarterbacks," *Res. Quart. Exercise Sport*, vol. A-72, 1993.

[23] R. A. Berk, Ed., *Performance Assessment*. Baltimore, MD: Johns Hopkins Univ. Press, 1986.

[24] H. Khan, "Procedural steps for designing and implementing an evaluative research using outcome based program evaluation," in *Frontiers in Education Conf.*, 1995.
[25] C. W. Anderson, K. M. Bryan, J. E. Froyd, D. L. Hatten, C. L. Kiaer, N. E. Moore, M. R. Mueller, E. A. Mottel, and J. F. Wagner, "Competency matrix assessment in an integrated, first-year curriculum in science, engineering and mathematics," in *Frontiers in Education Conf.*, 1996.
[26] R. W. Craine, D. C. Davis, D. E. Calkins, and K. Gentili, "Establishing engineering design competencies for freshman/sophomore students," in *Frontiers in Education Conf.*, 1995.
[27] A. D. Lucia, *The Art and Science of Competency Models: Pinpointing Critical Success Factors in Organizations*. San Francisco, CA: Jossey-Bass/Pfeiffer, 1999.
[28] L. H. Bradley, *Complete Guide to Competency-Based Education*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[29] G. O. Klemp, Ed., "Report to the National Institute of Education,", 1980.
[30] D. C. McClelland, "Testing for competence rather than intelligence," *Amer. Psychologist*, vol. 28, no. 1, pp. 1–14, 1973.
[31] F. Landy, S. Zedeck, and J. Cleveland, Eds., *Performance Measurement and Theory*. Hillsdale, NJ: Lawrence Erlbaum, 1983.
[32] *The Trainer's Handbook*, American Management Association, New York, 1998.
[33] L. Cohen, *Quality Function Deployment: How to Make QFD Work for You*. Reading, MA: Addison-Wesley, 1995.
[34] A. Bendell, J. Disney, and W. A. Pridmore, *Taguchi Methods: Applications in World Industry*. London, U.K.: IFS, 1988.
[35] W. E. Deming, *Quality, Productivity, and Competitive Position*. Cambridge, MA: Mass. Inst. Technol., Center for Advanced Engineering Study, 1982.
[36] G. Taguchi, *System of Experimental Design*: UNIPUB/Kraus Int., 1987.
[37] F. A. Kulacki and E. C. Vlachos, "Downsizing the curriculum: A proposed baccalaureate program and contextual basis," *J. Eng. Educ.*, pp. 225–234, July 1995.
[38] H. Khan, "Procedural steps for designing and implementing an evaluative research using outcome based program evaluation," *Proceedings*, vol. 1, pp. 2cl.9–2cl.12, 1995.
[39] C. W. Anderson, K. M. Bryan, J. E. Froyd, D. L. Hatten, C. L. Kiaer, N. E. Moore, M. R. Mueller, E. A. Mottel, and J. F. Wagner, "Competency matrix assessment in an integrated, first-year curriculum in science, engineering and mathematics," *Proceedings*, vol. 3, pp. 1276–1280, 1996.
[40] R. W. Craine, D. C. Davis, D. E. Calkins, and K. Gentili, "Establishing engineering design competencies for freshman/sophomore students," *Proceedings*, vol. 2, pp. 4d2.1–4d2.4, 1995.

William E. Dillon (S'64–M'68) received the B.S.E.E. degree from Texas A&M University, College Station, in 1965 and the M.S.E.E. and Ph.D. degrees from the University of Texas at Arlington in 1969 and 1972, respectively.

He is currently an Associate Professor of electrical engineering at the University of Texas at Arlington and is a Licensed Professional Engineer in the State of Texas. He has conducted research on power system modeling, solar power integration into the utility power grid, space power applications and high-voltage phenomena for EPRI, MIT Lincoln Laboratory, Sandia Corp., SDIO, BMDO, and LLNL. He has six years of industrial experience in the electric power industry. He has been active in curriculum development and student advising. In recent years, he has produced several interactive web courses.

George V. Kondraske (S'75–M'82–SM'90) received the B.S.E.E. degree from the University of Rochester, Rochester, NY, in 1978, and the M.S. and Ph.D. degrees in biomedical engineering from the University of Texas at Arlington/University of Texas Southwestern Medical Center at Dallas (joint program) in 1980 and 1982, respectively.

Currently, he is Professor of electrical and biomedical engineering and founding Director of the Human Performance Institute at the University of Texas at Arlington. He has conducted research sponsored by NSF, NASA, DOE, USAF, NIH, the Veterans' Administration, and the Texas Higher Education Coordinating Board. Key contributions include the development of GSPT, the Elemental Resource Model for human performance, and the Human Performance Capacity Measurement System. He has authored over 150 publications on these topics.

Louis J. Everett (M'88) received the B.S. degree from the University of Texas, El Paso, in 1978, the M.S. degree from Stanford University, Stanford, CA, in 1979, and the Ph.D. degree from Texas A&M University, College Station, in 1983.

Currently, he is an Associate Professor at Texas A&M and holds a joint appointment with the American University of Sharja where he is Associate Dean of Engineering. He has several years of industrial experience. He is a Member of the Management Team of the NSF Foundation Coalition and led the faculty teams that developed the revised undergraduate curriculum at Texas A&M. A key component of the NSF Foundation Coalition is outcome assessment and he has written several articles on the subject.

Richard A. Volz (M'60–SM'86–F'97) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Northwestern University, Evanston, IL, in 1960, 1961, and 1964, respectively.

He is Royce E. Wisenbaker Professor of engineering in the Computer Science Department at Texas A&M University, College Station. He was Department Head from 1988 to 1997, and has recently embarked upon a new initiative on "Training Systems Science and Technology." He has served on five federal advisory boards (Air Force Scientific Advisory Board, Ada Board, Aerospace Safety Advisory Panel—a Congressional oversight committee on NASA, NASA Space Station Advisory Panel, and NASA Center of Excellence in Information Technology Advisory Panel). He is the author of dozens of research papers, and has led numerous funded research projects.

Dr. Volz received the Decoration for Exceptional Civilian Service from the U.S. Air Force and two NASA Special Service Awards. He also served as Editor-in-Chief of the IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION.