AC 2010-1228: CONSTRUCTING MATHEMATICAL AND SPATIAL-REASONING MEASURES FOR ENGINEERING STUDENTS

Laura L. Pauley, Pennsylvania State University

Laura L. Pauley, Arthur L. Glenn Professor of Engineering Education and professor of mechanical engineering, joined the The Pennsylvania State University faculty in 1988. From 2000 to 2007, she served as the Professor-in-Charge of Undergraduate Programs in Mechanical and Nuclear Engineering. In 2003, Laura received the Penn State Undergraduate Program Leadership Award. Dr. Pauley teaches courses in the thermal sciences and conducts research in computational fluid mechanics and engineering education. She received degrees in mechanical engineering from University of Illinois (B.S. in 1984) and Stanford University (M.S. in 1985 and Ph.D. in 1988). She can be contacted at LPauley@psu.edu .

Jonna M. Kulikowich, Pennsylvania State University

Jonna M. Kulikowich, Professor of Education, joined The Pennsylvania State University faculty in 2003. Dr. Kulikowich teaches courses in applied psychometrics and statistics for the Department of Educational and School Psychology and Special Education. Currently, Dr. Kulikowich is an Associate Editor of the Journal of Educational Psychology. She is also Chair of the Faculty Senate Subcomittee on Undergraduate Education at Penn State. She received a psychology/philosophy degree from the University of Scranton in 1985 and degrees in educational psychology: research, measurement and statistics from Texas A&M University (M.S. in 1986 and Ph.D. in 1990). She can be contacted at jmk35@psu.edu.

Nell Sedransk, National Institute of Statistical Sciences

Dr. Nell Sedransk, Co-Pi, is the Associate Director of the National Institute of Statistical Sciences (NISS) and Professor of Statistics at North Carolina State University. She is an Elected Fellow of the International Statistical Institute, also Elected Fellow of the American Statistical Association. She has served as Associate Editor for the Journal of the American Statistical Association, the Journal of Statistical Planning and Inference, and has been Vice-Chair of the Publication Board of the American Statistical Association. The areas of her technical expertise and current research include design of complex experiments, Bayesian inference, spatial statistics and topological foundations for statistical theory. She received her Ph.D. in Statistics in 1969 from Iowa State University. She can be contacted at sedransk@niss.org

Renata Engel, Pennsylvania State University

Renata S. Engel is Associate Dean for Academic Programs and Professor of Engineering Design and Engineering Science & Mechanics. A member of the Penn State faculty since 1990, she served from 2000-2006 as the Executive Director of the Schreyer Institute for Teaching Excellence. Through various collaborative efforts, she has affected changes in the engineering curriculum at Penn State, primarily to incorporate elements of design in fundamental engineering courses. Engel earned a BS in engineering science at Penn State and PhD in engineering mechanics at the University of South Florida. She can be contacted at rse1@psu.edu.

Constructing Mathematical and Spatial-Reasoning Measures for Engineering Students

Abstract

Engineering students sometimes encounter difficulties in classes due to their ability to understand and interpret mathematical and visual representations of a problem. This paper describes tools to assess students' abilities in four different constructs. The two mathematical constructs are:

M1. Compare and contrast mathematical operations and

M2. Express engineering- and physics- based principles mathematically.

The two spatial-reasoning constructs are:

S1. Rotate and transform geometric objects in three-dimensional space and

S2. Translate two-dimensional images to three-dimensional images and vice-versa when representing visually engineering- or physics-based principles.

Examples are provided for each construct and assessment methods are also presented.

Background and Motivation

The purpose of this paper is to introduce mathematical and spatial-reasoning constructs that are keys to academic success in engineering. The term, "construct", is defined as a latent, unobservable trait, such as an ability or skill that directs how students select or generate answers to test items.[1] Several constructs or latent traits have been identified as important in engineering education. The authors illustrate how test items can be designed given various classroom assessment goals (e.g., course examinations, homework assignments) for two sets of constructs that can result in reliable and valid scores. Specifically, two mathematical constructs and two spatial-reasoning constructs are the focus of this paper. The mathematical constructs represent students' abilities to: (M1) compare and contrast mathematical operations (e.g., differentiation, integration, interpolation); and (M2) express engineering- and physics-based principles mathematically.

Likewise, two spatial-reasoning constructs are of interest. These constructs represent students' strategies to: (S1) rotate and transform geometric objects in three-dimensional space; and (S2) translate two-dimensional images to three-dimensional images and vice versa when representing visually engineering- or physics-based principles (e.g., acceleration, equilibrium, force).

Each mathematical and spatial-reasoning measure individually has received attention in the literature because of its importance in defining academic success in engineering. Devon, Engel, and Turner[2] determined that the students' ability to rotate and transform geometric objects in three-dimensional space is predictive of graduation and retention in engineering programs. Similarly, knowing how forces are represented visually in diagrams commonly employed in statics and thermodynamic courses is a skill that successful engineering students have. However, many college students have difficulty understanding how physics-based principles are represented visually. As a result, the types of problems assigned in courses like statics and thermodynamics that utilize these visual representations may be one reason these classes are perceived as challenging[3,4] and are sometimes called stumbling block courses.

The challenge students encounter in engineering courses is escalated by the fact that no ability or skill acts in isolation. Research from cognitive psychology[5,6,7] provides ample evidence that constructs must be coordinated or integrated if students are to reach levels of competence or proficiency within their domain. Therefore, in this paper, the researchers advocate for designing classroom measures that represent construct sets required to solve problems effectively in areas of specialization such as statics and thermodynamics.

The researchers also introduce psychometric questions to be addressed in the study concerning the reliability and validity of scores for the measures. These questions pertain to both dimensionality (i.e., how many constructs predict the response patterns for any given test) as well as how the scores are assigned for that test. Even for the well-known multiple-choice items, scores can be assigned in a variety of ways. For example, they can be scored dichotomously (i.e., correct versus incorrect) or polytomously (e.g., correct, partially correct, incorrect). Further, some of these multiple-choice items are constructed as testlets.[8] Testlets are groups of items that are dependent on the same stem or sets of tasks that must be solved within one problem space.[9] A reading comprehension passage followed by a set of multiple-choice items is a testlet. Determining the reliability and validity of testlet scores requires psychometric considerations that differ from those needed to analyze the data for a series of multiple-choice items that are independent and do not refer to a common stem or stimulus..

In the following sections, the investigators present items that represent each of the constructs of interest. The importance of these constructs within the coursework for engineering majors is described in the context of programs of study at one large university. Finally, a description of how the study of dimensionality and score assignment can lead to various reliability and validity analysis strategies is provided. Closing sections address statistical considerations and future directions in test and task development to study the academic development of students enrolled in undergraduate engineering programs.

Mathematical Test Items: Examples M1 and M2

The use of mathematics in solving and communicating engineering analysis can be an obstacle for some students. In describing the use of mathematics in engineering, we have distinguished between two different constructs, listed above as:

(M1) compare and contrast mathematical applications relevant to solving varied problems in engineering;

(M2) understand how the engineering quantities (e.g. force, work, power, and flow rate) are described by the mathematical representations (e.g. integration, differentiation, or interpolation) presented in statics, dynamics, thermodynamics, and fluid mechanics.

Although these two constructs are similar, we have listed them separately to better define the particular usage of mathematics that a student finds challenging. The following two examples will better define these constructs.

Construct (M1) refers to an understanding of the mathematical equations and solution methods without relating it to a physical quantity such as force, pressure, or power. An example of this type of problem is:

M1.7. The function $y = f(x)$ is shown on the graph. Circle all statements below that are true:

a. $\left.\frac{dy}{dx}\right|_1 > \left.\frac{dy}{dx}\right|_2$

b. $\left.\frac{dy}{dx}\right|_1 < \left.\frac{dy}{dx}\right|_2$

c. location 1 is an inflection point

d. $\left.\frac{d^2y}{dx^2}\right|_1 > 0$

e. $\left.\frac{d^2y}{dx^2}\right|_1 < 0$

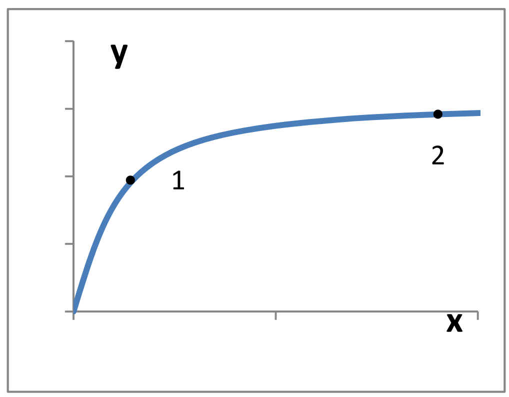


Figure 1. Example of Construct (M1).

To answer this question, a student must have an understanding of derivatives but there is no relation to physical quantities. Problems of this type can also be presented using different variables, say (*y*,*T*) instead of (*x*,*y*). Although the problem still uses variables with no physical interpretation, some students will find the second problem to be much more difficult because textbooks and instructors in calculus classes use (*x*,*y*) in most if not all problems. This finding might lead us to change the variable names throughout a calculus course and not always use (*x*,*y*).

The second use of mathematics tested is Construct (M2) that applies a physical meaning to the variables in the equation. An example of Construct (M2) is shown below. A second example can be found later in the paper as Figure 6.

M2.1. If h represents the height of water in a tank and t represents time, what does the following equation tell you about the height of the water in the tank?

$$\frac{dh}{dt} = -5$$

a. The height of the water is negative.
b. The height of the water does not change with time.
c. The height of the water is increasing with time.
d. The height of the water is decreasing with time.
e. Insufficient information given to answer the question.

Figure 2. Example of Construct (M2).

This question has added a physical meaning to each variable and asks for a physical interpretation of the differential equation. To answer this question requires several skills: understanding the definition of a derivative, using variables (t,h) instead of (x,y), and relating this equation to a physical process. For an unsteady problem such as this, the physical process cannot be easily communicated using a figure drawn on paper. Students must be able to mentally visualize a "movie" to understand the problem. Similar complications occur for three-dimensional problems that are shown as a two-dimensional representation.

Spatial-Reasoning Test Items: Examples S1 and S2

Two spatial-reasoning constructs are important in engineering education:

(S1) rotate and transform geometric objects in three-dimensional space; and

(S2) translate two-dimensional images to three-dimensional images and vice versa when representing visually engineering- or physics-based principles (e.g., acceleration, equilibrium, force).

Construct (S1) involves the ability to rotate and transform geometric objects in three-dimensional space. Similar to the Construct (M1) in mathematics, this spatial reasoning can be perceived as a general one that does not include reference to specific engineering- or physics-based principles. Yet, the literature documents clearly[2] that students who solve problems well in engineering have strong general spatial-reasoning strategies.

An example of Construct (S1) is shown below in Figure 3. This figure was used with permission from a Mental Rotation Test developed by Devon et al. [2, 10] A series of 12 rotation questions was developed to test the level and improvement of visualization skills in a freshmen design course. The researchers found that students' spatial visualization skills were improved more by using solid modeling than wireframe CAD or graphics taught in the traditional way in a freshmen design course.

S1.5 Which figure below is a rotation of the first?

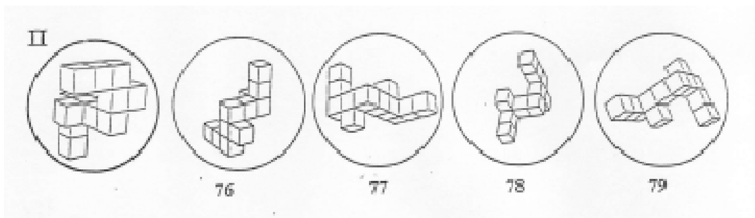


Figure 3. Example of 3D rotation, Construct (S1).

Construct (S2) requires translation of two-dimensional images to three-dimensional and vice versa when solving engineering problems. This construct includes the interpretation of figures, diagrams, and word descriptions that represent engineering- or physics-based principles. There are two different skills that are included in this construct:

a. Three-view two-dimensional projection drawing to a three-dimensional perspective drawing.
b. Relating different visual and mathematical representations of unseen quantities such as velocity, force, pressure, or temperature.

Engineering includes the analysis and interpretation of unseen quantities such as velocity, force, pressure, and temperature. Engineers often describe unseen quantities visually in graphs and figures. Students sometimes have difficulty in interpreting these graphs and figures, sometimes considering both coordinates as spatial coordinates and the plotted curve as a physical line or boundary. When the quantity is plotted using a cross-section of the geometry, the spatial visualization also presents a challenge. Below is an example of the laminar velocity profile in a pipe presented in three different ways: using a velocity profile, surface contour, and uniform velocity contours. Each representation includes two different answers. In each row of answers the student needs to decide if the first, the second, or neither of the figures describes the given velocity profile. These types of representations are used in many engineering courses. But students do not often admit that the figure is confusing. Questions about the figure are often realized when students ask questions about problem calculations. Are students hesitant to admit that the figure is confusing? Or does the student not realize that the confusion is in interpreting the figure?

S2.4. Mark all figures that are a visual representation of $u(r) = 1 - r^2$ where r is the radial coordinate and u is the velocity.

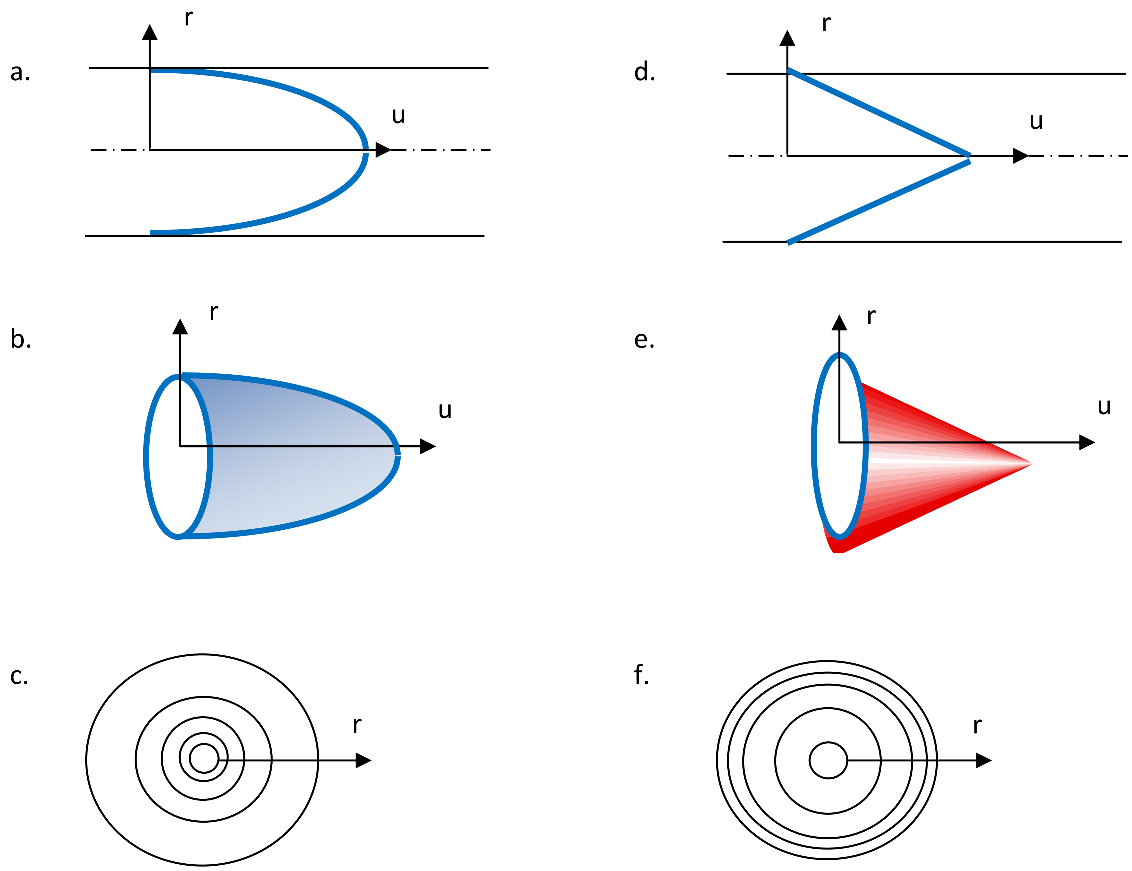


Figure 4. Example of various representations for Construct (S2).

<u>Test Design Strategy</u>

Design plans for all measures can include computer-based administration where response times for test completion can be recorded along with interactivity for certain tasks (e.g., manipulation of screen images). Currently, the investigators are planning pilot studies using paper administrations of the multiple-choice measures to acquire an initial examination of the reliability and validity of the scores. However, results from these studies will inform development of a computer-based administration tool. These tools can be programmed to produce so-called logfile[11] data that reveal not only the options or responses generated by students, but also, any changes in response, the time to complete any task or subtask, and use of notes, drawing, or highlighting to work through a problem space. Beyond information that can be scored as correct, partially correct, or incorrect, logfile data are important in that they can help instructors determine the sources of students' misconceptions or lack of prerequisite knowledge required to complete tasks successfully.[11,12]

The Importance of the Constructs in Programs of Engineering

Undergraduate programs in engineering share many characteristics. Usually, a prerequisite sequence of calculus, chemistry, and physics courses precedes declaring a major in several specializations. At Penn State, all engineering majors take an engineering graphics course in freshmen year. This course requires that students apply and strengthen their spatial-visualization skills (Construct S1). Students planning to pursue most engineering majors will then take a statics and a dynamics course in sophomore year. Most engineering students will also take a course in thermodynamics and/or fluid mechanics. These courses require students to draw and interpret visual realizations of non-visual quantities such as force, pressure, velocity, and acceleration. Often these quantities are visualized using vectors, profiles, or contours overlaid on the physical object(s). Since students have not seen this type of representation in the past, figures with physical geometry and non-visual quantities can be confusing. In addition, many calculus courses are taught with only x and y variables and some students have difficulty in applying the calculus methods to a new problem where derivatives of locations with time have the physical meaning of velocity and acceleration (Construct M2). Integration might be used in several different ways in engineering courses. In a thermodynamics course, integration of pressure and volume in a piston chamber gives work. In a statics course, students use integration to determine the center of gravity for the distributed weight of the object and the center of pressure for the distributed force on an object. If a student has difficulties in solving these problems, we need to identify the source of the difficulty. Is the difficulty in the understanding and solution of calculus (Construct M1), the spatial visualization and interpretation of the problem (Constructs S1 and S2), or the conversion of a complex problem to an equation form (Construct M2)? The test strategies described here attempt to dissect an engineering problem into these important steps to identify the source of the students' misunderstanding. Knowing the nature of the students' misunderstanding or misconception will allow for a focused solution to be developed.

While designing mathematical and spatial-reasoning measures informed by cognitive learning theory [11, 12] and understanding of multiple domains (e.g., engineering, mathematics, physics) is an important goal of classroom assessment practices, knowing how to assign scores reliably and validly given selections of responses or construction of them is essential in building a quality system of assessment that can provide important instructional information and feedback to advisers, students, and teachers. Historically, assignment of scores and address of their reliability and validity is a topic studied significantly in a specialization called psychometrics. [13] In the next section, the authors review various psychometric strategies to assign scores to items or tasks represented by constructs M1, M2, S1, and S2.

Psychometric Considerations

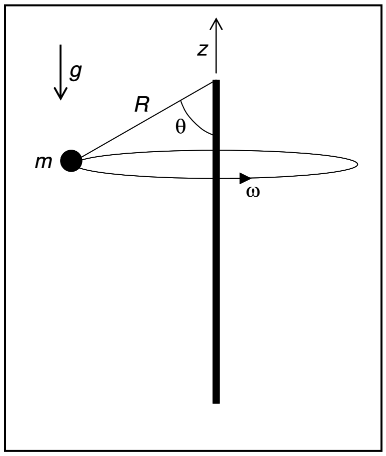
Sound measurement of latent variables depends on the reliability and validity of scores[14]. Reliability addresses how consistent the scores are over time while validity pertains to how well the scores are predicted by the constructs of interest. Two important decisions researchers must make to study reliability and validity concerns how many dimensions best represent the scores predicted by constructs or latent traits[1] and how scores are to be assigned given students' selection of options or construction of responses. Many constructs that involve complex reasoning and problem solving are multidimensional. As such, special modeling considerations are required to study score properties.

Related directly to assessment of dimensionality is the assignment of scores. Scores can be assigned dichotomously or polytomously. When scores are assigned dichotomously, responses are just measured in two categories as 1 (correct) or 0 (incorrect). Polytomous scoring assignment involves more than just two score categories (i.e., correct and incorrect). Additionally, there are a variety of polytomous-scoring assignments analysts can consider. These range from partial-credit scoring[15] to graded-response scoring[16] to scoring of response sequences or steps as in the case of procedural tasks, like solving for unknown values in algebra problems[17]. For this last category of scoring assignment, which requires evaluation of the paths or sequences of steps used by students, special psychometric modeling procedures are required to evaluate the conditional dependencies among responses. These psychometric modeling procedures usually are presented within a family of so-called item bundles, sets, or *testlets*.[8] Because both dichotomous and polytomous-scoring with and without testlet design strategies are important given the investigators' program of research, they are described with examples in this article.

Example One: Dichotomous Scoring with No Testlet Design Strategy

The most common of all achievement or aptitude item formats remains the multiple-choice item. Not only are multiple-choice items easy to administer, but they are also easy to score. Further, their scores can be analyzed quickly for reliability and validity; particularly in the instance of dichotomous scoring where the item is not part of a testlet. Figure 5 is a multiple-choice item on one of the spatial-reasoning measures (i.e., Construct S2) where the selection of the response is scored correctly or incorrectly. Further, there is only one item that corresponds to the stem and graphical stimulus, so this item is not part of a testlet.

S2.1. A rigid object of mass *m* attached to a string of length *R* rotates in a circle at constant angular speed ω about a pole, as shown. Gravity acts parallel to the axis of rotation, as indicated. The force exerted by the string on the mass is directed along the axis of the string.



What is the direction of the instantaneous acceleration of mass *m*?

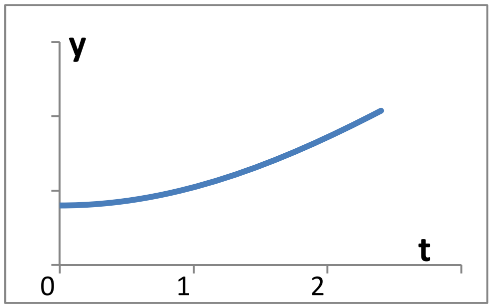
a. -*z* direction
b. +*z* direction
c. Parallel to the string
d. Radially inward towards the pole
e. Radially outward from the pole

Figure 5. Example of dichotomous question for Construct (S2).

Example Two: Assigning Scores with a Testlet Design Strategy

By comparison, Figure 6 displays an item for the measurement of physical principles as they are represented mathematically (i.e., Construct M2). This item is a testlet. Upon review, the item includes a graphical stimulus in the stem similar to that of the spatial reasoning item. However, several options treated as separate items that reference the stem may be considered correct. For example, the options reference three time points depicted in the graph, thus all statements pertain to the same function. At time point 0, both options a and b would be correct responses. However, at times points 1 and 2, there is only one correct selection. As such, all other selections, relative to each time point, would be false. Analysis of scores requires careful examination of correct and incorrect response patterns. These patterns could be examined for each time point, and then statistically, the time point responses are nested within the pattern for the complete item, sometimes called the superitem.[17]

M2.8. The location of a car y(t) is a function of time t as shown in the graph.



Circle all of the statements that are true:

a. At t=0 the car is stopped and begins to move.
b. At t=0 the car is accelerating.
c. At t=0 the car is traveling at a steady speed.
d. At t=0 the car is decelerating.
e. At t=1 the car is accelerating.
f. At t=1 the car is traveling at a steady speed.
g. At t=1 the car is decelerating.
h. At t=2 the car is accelerating.
i. At t=2 the car is traveling at a steady speed.
j. At t=2 the car is decelerating.

Figure 6. Example of a testlet for Construct (M2).

The Psychometric Evaluation of Dimensionality and Score Assignment

Typically, classical exploratory and confirmatory factor analytic and Item Response Theory methods are used to determine the dimensionality and precision of score assignment.[13,14] These techniques are useful primarily when each distribution represents a single dimension and scores are assigned dichotomously. For several of the measures administered in our research (e.g., mathematical or spatial reasoning), scores are likely to reflect multidimensionality. For instance, Construct (M2) designed to measure understanding of mathematical information includes as much scientific content as mathematical information. From a theoretical perspective, therefore, these two domain sources of information must be integrated to answer the items correctly. Further, students with varying levels of ability may not integrate the sources to similar degree leading to response patterns that reflect both lack of mathematical and scientific knowledge, and the lack of knowledge by domain may not be similar in magnitude leading to what psychometricians refer to as mixtures. [18, 19] Mixture modeling has historically been applied when multidimensionality is assumed. Because testlets often require students to process various sources of information during problem-solving, many models required to study score patterns for dependent responses also test for multidimensionality. In the current program of research, the investigators intend to study the benefits of applying various IRT models to establish that the psychometric properties of scores are sound for the various mathematical and spatial-reasoning constructs (i.e., M1, M2, S1, S2). In the final two sections, directions for future research in statistical planning and item development are summarized.

Statistical Planning Considerations

Establishing the dimensionality and precision of score assignment for the four constructs of interest is essential for statistical planning. Currently, our research team is focusing on three inter-related projects. The first is a planned statistical experiment to determine the degree to which apprenticeship opportunities with Engineering Design Principles increases scores on the measures. If the scores are not reliable and valid as outcomes, then test of intervention efficacy and estimates of effect size are not possible.[18] Further, planning of statistical experiments requires careful consideration of how many independent and dependent variables are to be included in any model that is tested. The address of dimensionality informs directly the frequency of scores to be included in a statistical model.

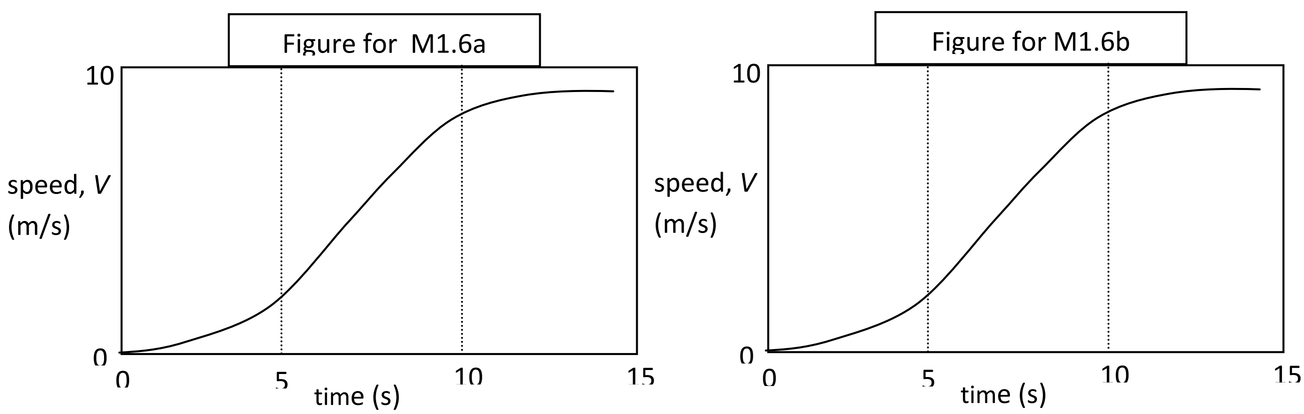
Reliable and valid scores of the measures are also to complement survey research initiatives. Currently, the researchers are studying how attitudes and levels of engagement are related to academic performance. Studies in academic development [5,11] demonstrate that it is not only important to look at achievement-type variables such as reading comprehension, mathematical reasoning, and problem solving, but also, understanding of affect-type variables like interest and motivation provides opportunity to identify experiences that are optimal learning opportunities for students.

Finally, accreditation is important in programs of engineering. Reports that document steps taken to improve curriculum development, advising, and teaching effectiveness need to rely on data, which highlight the strengths of the program and identify areas where improvement can be made. Test scores that reflect the key constructs in Engineering courses, such as those described in this article, which are reliable and valid add to the quality of any systematic evaluation.

Future Research Directions on Test Development

There are many current developments in the program of assessment research described in this article. In addition to administering measures using a computer tool, plans include the design items that allow students to construct or to generate their responses. Figure 7 presents an example of an item where students are afforded more opportunity to interact with the stimuli. Specifically, this constructed-response testlet requires that students draw to demonstrate their knowledge of derivatives and integrals for a task that is sampled to represent Construct (M1). With advances in computer programming, and as mentioned previously, students' generated responses can be captured not only as images, but also, any changes or revision they make when constructing their responses as well as the time it takes for them to complete each task can be recorded by the computer system. While analysis of constructed responses can introduce complexities when examining both the dimensionality and precision of scores, there are advantages. One advantage in particular, and unlike that observed in many multiple-choice responses, is that guessing is reduced. This approach can be extended to Constructs of type M2. For example Figure 7 tasks could be to "draw something that shows qualitatively the acceleration at time t=10 sec" and 'to draw something that shows qualitatively the distance traveled during the time interval from 5 to 10 seconds." Further, the opportunity to follow the processing of students as they construct their responses may provide important insights as to misconceptions they have. These misconceptions may be difficult to detect when relying solely on multiple-choice formats.

M1.6. In the figures below we plot the speed of a car, *V*, vs. time, *t*.



(a) On the left figure draw something on the plot that shows qualitatively what is the derivative of *V* with respect to *t* at the time t=10 s.

(b) On the right figure draw something on the plot that shows qualitatively what is the integral of *V* over time from 5 to 10 seconds

---

Figure 7. Example of open answer question for Construct (M1).

Conclusions and Significance

Reliable and valid test scores are needed for domain-specific measures in engineering to not only profile patterns of strength and weakness for students enrolled in various programs, but also to test instructional interventions that may facilitate academic progress. This presentation introduces test design strategies and data-analytic considerations in the study of the psychometric evaluation of scores of two mathematical and two spatial-reasoning constructs considered key to academic success in engineering.

References

[1]Borsboom, D, (2005), *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge, UK: Cambridge University Press.

[2]Devon, R., Engel, R., Turner, G. (1998), "The Effects of Spatial Visualization Skill Training on Gender and Retention in Engineering," *Journal of Women and Minorities in Science and Engineering* **4:** 371-380.

[3] Dollár, A., Steif, P. S. (2006), "Learning modules for statics," *International Journal of Engineering Education* ***22(2):*** 381-392. Retrieved from http://www.ijee.dit.ie/ OnlinePapers/Interactive/Statics/Learning-Modules-for-Statics-Dollar-Steif.html

[4] Huang, M., Gramoll, K. (2004). "Online interactive multimedia for engineering thermodynamics," *Proceedings of American Society for Engineering Education Annual Conference & Exposition*. Retrieved from http://soa.asee.org/ paper/conference/paper-view.cfm?id=20564

[5]Alexander, P.A. (2003), "The development of expertise: The journey from acclimation to proficiency," *Educational Researcher* **32**: 10-14.

[6]Chi, M. T. H., Feltovich, P. J., Glaser, R. (1981), "Categorization and representation of physics problems by experts and novices," *Cognitive Science* **5**: 121-152.

[7] Linn, M. C. (2000), "Designing the knowledge integration environment," *International Journal of Science Education* ***22*****(8):** 781-796.

[8]Wainer, H., Kiely, G, (1987), "Item clusters and computerized adaptive testing: A case for testlets," *Journal of Educational Measurement* **24**: 185-201.

[9]Almond, R. G., Mulder, J., Hemat, L. A., Yan, D, (2009), "Bayesian network models for local dependence among observable outcome variables," *Journal of Educational and Behavioral Statistics* **34(4)**: 491-521.

[10]Devon, R., R.S. Engel, R.J. Foster, D. Sathianathan, G.F.W. Turner, (1994), "The Effect of Solid Modeling Software on 3-D Visualization Skills," *Engineering Design Graphics Journal* **58(2):**4-11.

[11]Alexander, P. A., Murphy, P. K., Kulikowich, J. M. (2009), "Expertise and the adult learner," In M. C. Smith (Ed.), *Handbook of research on adult learning and development* (pp. 484-523). New York: Routledge, Taylor and Francis Group.

[12]Kulikowich, J. M. (2007), "Toward developmental trajectories: A commentary on assessing measures of mathematical knowledge for teaching," *Measurement: Interdisciplinary Research and Perspectives* **5(2-3)**: 187-194.

[13]Lord, F.M., Novick, M.R., (1968), *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

[14]McDonald, R.P. (1999), *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

[15]Masters, G.N. (1982), "A Rasch model for partial credit scoring," *Psychometrika* **47**: 149-174.

[16]Samejima, F. (1969), "Estimation of ability using a response pattern of graded scores," *Psychometrika Monograph*.

[17]Mislevy, R. J., Verhelst, N. (1990), "Modeling item responses when different persons employ different solution strategies," *Psychometrika* **55**: 195-215.

[18]Kulikowich, J.M., Sedransk, N. (2010), "Current and emergent design and data analysis approaches" Manuscript in preparation.