April 3, 2013

To: Mark Greenberg

President

 New England Tractor Trailer Training School

From: Ellen Wentland, Ph.D.

Re: Determining the Effectiveness of Entry-Level Commercial Driver Training Programs

**Introduction**

Recent years have seen an increased emphasis on the part of consumers, supporters, employers, and oversight groups on accountability and transparency in the delivery of educational preparation and career training programs. This occurred at least in part because of perceptions that the content promised or said to be delivered by these programs was often not realized in terms of what graduates knew or were able to do. The simple specification of number of credit hours or hours spent on topics was not sufficient to assure the covering of program content or, more importantly, adequate mastery on the part of students of that content.

Even when program content is carefully specified, variations in factors such as instructors’ interpretation of content, methods of delivery, and areas of emphasis as well as in students’ prior preparation and receptivity contribute to a learning experience which is typically not uniform within and across classes or schools for the students involved. Beyond specifying content and number of “exposure to content” hours, what is needed to support program quality is a set of expectations for what successful students will learn and methods to evaluate students against these expectations. Concern about program quality has led to a national change of focus from hours spent on topics, or student “seat time”, to an emphasis on determining what students have acquired in the form of knowledge and skills. The extent to which students achieve specified learning outcomes or competencies has become the measure of program effectiveness.

**Importance of Knowledge and Skills Assessments**

The new emphasis is accompanied by increased demands for and use of knowledge and skills assessments within educational preparation and career training programs. These assessments are tools that demand a close attention not only to content, but to outcomes and the impact of program educational experiences on the students’ acquisition of knowledge and skills. This type of rigorous attention to program effectiveness is important across the board, but may be especially critical in occupations and professions that involve high levels of responsibility or risk to the worker and /or the public, for example, in the areas of health or in occupations that require the use of hazardous substances or heavy equipment or machinery.

Programs in the health professions typically include highly specific and prescribed program content aligned with state and/or national guidelines, descriptions of expected learning outcomes, and methods of student evaluation with respect to these outcomes. Even in cases where program graduates are subject to governmental licensure requirements for employment, extensive knowledge and skills assessments are conducted within the various professional and career programs to assure and document program quality and effectiveness, namely, that students have developed the specific required competencies.

Students in Radiological Technology programs, for example, are provided with a variety of classroom, practical, and clinical experiences designed to support students’ achievement of the learning outcomes necessary for the profession. In addition, to assure a satisfactory level of knowledge and skills acquisition, students are repeatedly tested and evaluated. Eligibility to sit for licensure examinations can be contingent on documented evidence from these programs supporting students’ satisfactory demonstrations of competencies with respect to all components of the program curriculum.

As another example, students preparing for certification as Emergency Medical Technicians are not only provided with a curriculum, but are also required to take knowledge and practical skills examinations within the course of the program to ensure adequate mastery of required academic and clinical knowledge and skills. Unsatisfactory performance on the tests can result in students being required to carry out additional educational activities and then to re-take examinations, or in the case of multiple failures, to repeat the program.

An example of a risk-associated occupation which requires complex performance skills is crane operation. Applicants for certification must pass a set of written and practical skills tests, and certain states - West Virginia and New Jersey - now require this certification as a condition of eligibility for taking the state licensing examination.

**Knowledge and Skills Assessments of Students in Entry-Level Commercial Vehicle Training Programs**

Drivers of commercial vehicles are also engaged in an occupation that involves risk and responsibility to the public. The knowledge required is quite extensive, and covers a wide number of areas, including regarding federal and state highway regulations, general safe operating practices, and vehicle maintenance. Considerable performance skills are also required, such as regards vehicle inspection and operation. While entry-level commercial vehicle driver training programs exist across the country, there appears to be consensus concerning the general content needed to be covered for adequate and comprehensive preparation. Because content “coverage” does not assure student learning, however, a necessary next step is knowledge and skills assessments.

Becoming an entry-level commercial driver requires passing the state Commercial Driver’s License (CDL) test. This test, however, is limited in time and in its coverage of critical topics. As such, it may be inadequate to serve the purpose of ensuring satisfactory mastery of the extensive set of knowledge and skills necessary for effective driving performance. This purpose can be better served through the use of competency assessments at the individual program level.

These program-level assessments may provide a more comprehensive focus than the CDL test on the relevant content areas related to successful preparation for entry-level commercial drivers. These type assessments can provide systematic feedback allowing for timely corrections that minimize potential negative outcomes for students and the public. For example, at the school level, the assessments can be used to identify areas of individual student weakness, allowing students to re-take certain content prior to program advancement. And should any patterns of weakness emerge across students, improvements to pedagogy or resources in the form of, for example, equipment and/or textbooks can be made.

Currently, little information is available regarding the level of preparedness of individuals for commercial driving - the level of competency acquired - following the completion of the various entry-level commercial driver training programs. Knowing the length of a program, or the amount of time or hours spent on various topics, does not provide information on what students know or are able to do once they complete the program. Further, without knowing what students have learned, it is not possible to identify “best practices” in terms of time and methods. In general, absent some set of common learning outcomes and cross-program assessments, little is known about the relative quality of the training being provided in these programs.

**Development of a Valid Assessment of Students in Entry-Level Commercial Vehicle Training Programs**

Tests to validly assess student competency acquired at individual programs can be constructed with the use of a set of well-established test development models and procedures, necessarily including input from subject matter experts. An important first step would be the description of required program content, through the specification of topic areas. Next, the topic areas would be expanded to a degree of detail in the form of learning outcomes – what successful students will know or be able to do - that link to knowledge items or questions for a written test, or performance tasks for skills tests. These tests would be accompanied by scoring guides for written tests, detailed rubrics for performance tests, instructions on the application of the performance rubrics, and the establishment of test performance standards.

As an example, such a process was followed in the recent development of the **Assessment of Entry-Level Commercial Drivers (AECD)**, a knowledge and skills assessment for students completing commercial vehicle driver training programs. This assessment includes two versions of a 125-item multiple choice written test, yard performance tests, and a road test.

**General Steps**

Very briefly, the general steps followed included:

* Identifying twenty-four topics in five major subject areas that were judged by subject matter experts to be required for entry-level commercial driver training.
* Determining the method of assessment for these topics; written, yard and/or road.
* Identifying Evidence Centered Design (ECD) (Zieky, 2006) as the model that would guide the test development efforts.
* Outlining the topics identified to be assessed in accordance with ECD model.
* Specifying items and tasks in a way as to provide specific linkages between what we want to say about test-takers concerning their learning and the activities (answering items or performing tasks) that would provide the supporting evidence.

The validity of the test content was achieved through the linking of the written test items and performance tasks to the content areas (Fuhrman, 1996). The content was carefully delineated, at general and then increasingly specific levels, through consultation with subject matter texts and experts in the field. Items and tasks were then developed to link or relate to the identified content.

**Development of the Written Tests**

The development of the two versions of the written test included:

* Following rules developed by experts for writing multiple-choice test items (e.g. Fuhrman, 1996; Gross, 1994; Haladyna, 1999; Haladyna and Downing, 1989; Linn, 1995; and Osterlind, 1998).
* For content, using expert assistance in the form of textbooks on truck driver training widely accepted in the training industry, federal regulations, and extensive input and reviews by a panel of subject matter experts in the field.
* Conducting a series of pilot tests of the two test versions, which involved:
* Developing and distributing a detailed curriculum based on the topic areas outlined.
* To ensure test integrity, security, and standardization of administration, providing detailed instructions to school test coordinators as well as to test administrators.
* Again to ensure test integrity, requiring students and test administrators to attest to the following of test procedures and instructions through providing their signatures.
* To assure test item quality, conducting various item analyses identified by experts in the field (e.g. Ebel and Frisbie, 1991; Linn, 1995; Osterlind, 1998; and Thorndike, 1971).

**Development of the Performance Tests – Yard and Road**

The development of the yard and road performance tests also included the extensive involvement of subject matter experts and involved:

* Identifying specific topic-related tasks, with expert verification of the relatedness of the tasks to the defined content.
* To help eliminate subjectivity from ratings, creating Scoring Guides containing detailed task descriptions and performance standards.
* To ensure standardization of administration, creating rating forms with the instructions that are to be given to students.
* Conducting a series of pilot tests including:
	+ Providing detailed instructions to schools regarding general performance test administration.
	+ Specifying procedures to estimate reliability of yard and road test ratings

**Final Steps**

Final steps, again with the extensive involvement of subject matter experts, included the development of standards for both the written and performance tests. Also, detailed score reports were created which included information by topic area not only on individual student performance, but also on group level performance vis a vis the agreed upon performance standards. Clearly, this type of rich feedback is quite valuable for schools as they can celebrate successes, identify specific areas for individual or group improvements, and monitor effects of improvement efforts over time.

**School-Based Test Administration and Test Security**

Concerning test administration, approaches commonly used to assure fairness and standardization include having regional testing centers – the model used by state departments of motor vehicles and by administrators of the ACT and SAT as well as numerous testing organizations. Another model is the school-based one used, for example, for the administration of the PSAT. In that model, authorized individuals at schools are provided with extensive training regarding test administration, for example, in the form of detailed guides and instructions. Test security and integrity is addressed through a number of procedures such as providing detailed instructions regarding test storage and access, requiring students as well as instructors to attest to the integrity of the test administration and test-taking, and using the services of a testing organizations to distribute, collect, continuously monitor, and refine the testing instruments.

**School-Based Test Administration in Entry-Level Commercial Vehicle Driver Training Programs**

Because of the large number of programs across the country providing entry-level commercial driver training, the large number of students involved - approximately 70,00 to 100,000 each year, the staggered nature of program start and end times at the various schools, the sequencing of classroom and yard/road instruction, and the need for extensive and individualized yard and road testing for students in the commercial vehicles, the latter school or program-based test administration model would be the most appropriate and feasible, and was the one adopted for the administration of the AECD.

**Recommendation for Involvement of a Testing Organization to Ensure Test Security, Integrity and Validity**

In addition to the procedures described above and used in the AECD administration and other school-based testing models, the assistance of a testing organization to assure ongoing test integrity, security, and validity of test content is recommended.

The work of a central testing organization could include:

* Training and certifying test administrators to ensure test integrity and standardization of procedures.
* Performing various data forensics and security processes such as related to unexpectedly high or low school pass rates.
* Generating topic specific score reports at the individual and school levels, to allow for examination of student performance at points in time and over time, and for the identification of specific areas in need of improvement.
* Delivering test administration training and materials.
* Conducting various test item analytics to examine test item quality, including checks on item difficulty levels and test reliability.
* Examining patterns of test item responding across individuals tested in the same classroom or across classrooms for detection of any irregularities.
* Implementing an incidence reporting system for cases of irregularities.
* Developing additional test questions to form a test item pool that can be used to create multiple test versions.
* Developing new items or tasks to reflect changes in applicable laws, technologies, or the vehicles.
* Conducting continuing reliability checks on performance test ratings.
* Conducting periodic site observations.

**Summary and Conclusions**

Information concerning the quality or effectiveness of entry-level commercial vehicle driver training programs can be obtained through the use of well-constructed school-based knowledge and performance tests, developed and administered in accordance with professional testing principles. The Assessment of Entry-Level Commercial Drivers (AECD) is an example of such a test that can be used within programs across the country. Unlike an approach which focuses exclusively on time spent on content, the assessment approach allows for the systematic collection of extensive information regarding individual and school performance by specific topic areas. This approach provides ready feedback which allows for corrections in the form of requiring students to review content before advancement or examining methods to improve overall student content mastery. Importantly, this approach help ensure that students participating in these programs are receiving the knowledge and skills training at a level necessary for safe and effective performance as a commercial vehicle driver.

*Ellen Wentland, Ph.D., is Associate Dean of Academic and Institutional Assessment at Northern Essex Community College. She has a doctorate in Social and Organizational Psychology from the University of Chicago, and has taught courses at the graduate level in educational and psychological tests. Further, she has worked closely with test development specialists and statisticians at The College Board and Educational Testing Services in exploring and implementing changes to the PSAT and SAT. At the request of representatives for the Commercial Driver Training Foundation, and with the assistance of subject matter experts from across the country, in 2011 Dr. Wentland participated in developing the* ***Assessment of Entry-Level Commercial Drivers (AECD)****, a knowledge and skills assessment of students in entry-level truck driver training programs.*

**Bibliography**

Ebel, R.L., & Frisbie, D.A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Fuhrman, M. (1996). Developing good multiple-choice tests and test questions. *Journal of Geoscience Education, 44,* 379-384

Gross, L.J. (1994). Logical versus empirical guidelines for writing test items. *Evaluation and the Health Professions, 17*(1), 123-126.

Haladyna, T.M. (1999) *Developing and validating multiple-choice test items.* Mahwah, NJ: L. Erlbaum Associates.

Haladyna, T.M., & Downing, S.M. (1989a). A taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education, 1,* 37-50.

Linn, R.L., & Gronlund, N.E. (1995). *Measurement and assessment in teaching* (7th edition). Englewood Cliffs, NJ: Prentice Hall.

Osterlind, S.J. (1997). *Constructing test items* (2nd edition). Boston, MA: Kluwer Academic Publishers.

Thorndike, R.L. (1971). *Educational measurement* (2nd ed.). Washington, D.C.: American Council on Education.

Zieky, M. (2006). An introduction to evidence centered design. Princeton, NJ: Educational Testing Service.