

Chapter 16

Learning Performance in Vocational Secondary Schools: Testing Academic Achievement in Electrical Engineering



Metwaly Mabed and Thomas Köhler

16.1 Introduction

16.1.1 *Learning Performance, Academic Achievement, and Knowledge*

Learning performance links to the future career opportunities, especially, in the vocational education field (Adams 2014). Therefore, it is not surprising that the academic achievement of students is a top concern of educators. Academic achievement is often used to refer to the knowledge obtained by students through a school program or curriculum. According to Algarabel and Dasi (2001), academic achievement is defined as “the competence of a person in relation to a domain of knowledge” (p. 46) or the proficiency of students’ performance in a certain course. From the given importance of this topic in an educational context, it is necessary to do the process of exploring, assessing, and evaluating academic achievement by a particular test. AERA, APA, and NCME (AERA et al. 1999) pointed out that academic achievement tests “are measures of academic knowledge and skills that a person acquired in formal and informal learning opportunities” (p. 124). Other definition of academic achievement test can be found in the *Dictionary of Education*, which defines “a test that measures the extent to which a person has acquired certain information or mastered certain skills, usually as a result of planned instruction or training” (Shukla 2005). Consequently, academic achievement tests tend to measure

M. Mabed (✉)

Faculty of Education, Suez University, Suez, Egypt
e-mail: metwalys@gmail.com

T. Köhler (✉)

Faculty of Education, Technische Universität Dresden, Dresden, Germany
e-mail: thomas.koehler@tu-dresden.de

© The Author(s) 2018

J. Drummer et al. (eds.), *Vocational Teacher Education in Central Asia*,
Technical and Vocational Education and Training: Issues, Concerns
and Prospects 28, https://doi.org/10.1007/978-3-319-73093-6_16

151

recent learning performance and are closely tied to particular subjects or courses (Ariyo 2007). As outcome of a doctoral research and development project, the presented test has been discussed by several reviews. The presented publication results from a conference presentation at KICSS 2014 (Köhler and Mabed 2014; Mabed and Köhler 2016), focusing on the test construction and its quality. The recent publication intends to provide access to the test itself in a way that it may be reused easily in regular educational situations.

16.1.2 Testing Academic Achievement

Testing academic achievement may serve three major functions (Srivastav 2000): (a) measuring the effectiveness of learning performance at a particular point of time in an educational program; (b) predicting the behavior of students across different situations; and (c) assessing various psychological traits and characteristics. The information, which is collected by the achievement test, moreover, supports the decisions concerning the placement or diagnostic issues during the instruction process. Furthermore, Smith (2001) argues that there is a wide range of reasons of using academic achievement tests. Those are to (a) identify where the student falls along continuum of knowledge acquisition; (b) classify students in groups according to their scores in the end of a course; (c) determine the eligibility for particular educational programs; and (d) measure the effectiveness of instruction process. Whether the achievement test is a norm-referenced measure or a criterion-referenced measure, the vital issue is that its results should represent accurately the evidences concerning student performance to enhance the learning process as a central goal in any educational organization.

In vocational education, indeed, the precise specification of what to measure is poorly understood. Only recently, a renewed understanding of how testing learning outcome is linked to teaching activity is taken into account for developing measure. The development of academic achievement tests in order to assess students' learning performance in electrical engineering is more beneficial for research and practice. Although many researchers are interested in measuring academic achievement (Bayrak and Bayram 2010; Carle et al. 2009; Choy et al. 2012; Haislett and Hafer 1990; Rivkin et al. 2005; Romano et al. 2005; Selçuk et al. 2011), no existing measure was found to be adequate to deal with the electrical engineering content in this study. The necessity for the construction of such a test cannot be understood without discussing the influence of an achievement factor in many aspects of students' professional skills in vocational secondary school. Therefore, the construction and development of an appropriate test to measure students' achievement in electrical engineering will enhance the quality of learning outcomes.

16.2 Empirical Method of Test Development and Resulting

16.2.1 Aims, Procedures, and Data Analysis

The aim of this study is twofold. The first purpose is to identify the learning performance that reflects students' assimilation of the electrical engineering content. The second purpose is to describe the construction of a reliable and valid academic achievement test designed to measure students' learning performance. To achieve these goals, the researchers developed an academic achievement test in several phases by following the systematic approach and the procedures, which have been provided in Crocker and Algina (1986), Downing (2006), and Osterlind (1998) as well as in the light of the standards (AERA et al. 1999). Statistical analysis was conducted using PASW (Predictive Analytics SoftWare) Version 18.0 for Windows.

16.2.2 Development Phases

The construction of the academic achievement test consisted of eight subsequent development phases as follows:

1. Clarifying the purpose of the academic achievement test (Crocker and Algina 1986)
2. Identifying the educational objectives (Bloom et al. 1956; Crocker and Algina 1986)
3. Performing panel of experts
4. Developing test blueprint (AERA et al. 1999; Crocker and Algina 1986; Osterlind 1998)
5. Determining and generating test questions (Osterlind 1998; Quellmalz and Hoskyn 1995)
6. Preparing test instructions (AERA et al. 1999)
7. Performing panel of experts (Crocker and Algina 1986) and
8. Conducting the pilot study.

16.2.3 Resulting Test

The main outcome of the presented study is the test to be applied in the context of regular vocational or general school activity. Due to the field condition, it has been developed in both English and Arab versions. The complete achievement in electrical engineering test may be found in the annex and was as well published in the dissertation of Mabed (2013).

The test in its English version starts with an easy to understand test instruction to be read by the students before completing it. The main section of this test consists of 60 multiple questions. Each question has four possible answers (A, B, C, and D), and one must choose the correct or best answer. As well, a pocket calculator is recommended for students when completing this test. Answers to all questions are collected on a single final sheet with answer boxes for each question. Afterward this sheet may be easily used by the teacher for calculating the outcome on the basis of a graphical template.

Even though the test would be available as a browser-based version, it is designed in a way to work as paper and pencil as well, without using a computer in order to conduct the test, which does not necessarily need to take place online. The duration of the test has suggested with approximately 100 min in order to provide an adequate time for answering the whole test with all its items (cf. Mabed 2013) and should be applied in the context of regular educational activities in vocational secondary schools.

16.3 Findings on Test Quality

16.3.1 Test Validity

The researchers were keen to utilize a valid academic achievement test. Therefore, it was a focus to check the validity of academic achievement test. However, test validity is defined as “the process of collecting evidence to establish that the inferences, which are based on the test results, are appropriate” (Adams 2014, p. 21). From this meaning, the validity term is related to the scores, which are obtained from applying the test more than the test items itself. In the same stream, a complete description of the validity issue is found in (AERA et al. 1999): “Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing and evaluating tests. The process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations” (ibid., 1999, p. 9). The process of assessing validity includes many actions such as asking test takers and experts about their subjective opinions on the content of the test items, as well as using factor analysis to evaluate the item internal structural (Kline 2005). In this study, the principal components factor analysis with a varimax rotation statistical procedure was employed to calculate the factor loadings and eigenvalues of the answers for the sample by using the statistical program PASW Version 18.0.

Two criteria were determined to achieve satisfactory validity of the test items. First, the factor loadings should exceed the minimum threshold of 0.5 as recommended (Hulland 1999). Second, the items with multiple loadings should record a higher load on their related level than the load value on other levels (Gefen and

Straub 2005). The results revealed that the factor loadings of the five items were less than the cutoff of 0.5. Moreover, four items were loaded greater onto another level than their corresponding level.

In summary, 9 out of the 70 items were removed from the academic achievement test. Then, the factor analysis with the varimax rotation was run again over the dataset for 61 questions. The results showed that the factor loading value of each item was at 0.5 or above and there was no cross loadings in between the test items. The first component referred to the application level, and it had an eigenvalue of 20.89. The second component indicated to the knowledge level, and it had eigenvalue of 9.86, while the third component referred to the analysis level, and it had eigenvalue of 5.00. Finally, the fourth component pointed out to the comprehension level, and it had eigenvalue of 3.22. Accordingly, all eigenvalues were more than the cutoff value of one. Moreover, the four levels accounted 64% of the total variance.

16.3.2 Test Reliability

Reliability as an indicator was calculated during the construction of the test. While the test was defined as a scale describing student behavior in a specified domain, test reliability refers to the stability of this scale when the testing procedure is repeated on a population of individuals or groups (AERA et al. 1999). Cronbach's alpha was 0.97 higher, that is than the restrictive criterion of 0.7 (Fornell and Larcker 1981). Since the academic achievement test consisted of four levels (knowledge, comprehension, application, and analysis), it is not sufficient to compute the reliability for only the entire test when the score of sublevel will be used. According to standard 2.1 (AERA et al. 1999, p. 31), "For each total score, sub-score, or combination of scores that is to be interpreted, estimates of relevant reliabilities and standard errors of measurement or test information functions should be reported." Thus, the reliability for each level in the achievement test was also reported, as shown in Table 16.1. In regard to the knowledge level, Cronbach's alpha was 0.967 for the 15 items, and the item correlation values ranged from 0.69 to 0.88.

Concerning the comprehension level, the reliability coefficient alpha Cronbach was 0.94 for the 15 items, and the item correlation values ranged from 0.51 to 0.77. In regard to the application level, an internal consistency analysis showed alpha value of 0.96 for the 22 items, and the item correlation values ranged from 0.52 to 0.86. Regarding to the analysis level, Cronbach's alpha coefficient was 0.94 for the

Table 16.1 Results of knowledge level reliability on the academic achievement test

Test level	No. of items	Cronbach's alpha	Mean	Variance	SD
Knowledge	15	0.967	11.42	28.22	5.31
Comprehension	15	0.939	9.71	26.71	5.17
Application	22	0.958	14.68	56.43	7.51
Analysis	9	0.936	4.54	13.14	3.62

nine items, and the item correlation values ranged from 0.53 to 0.92. As a result, the items of levels in the achievement test presented satisfaction reliability criteria.

16.3.3 Item Analysis

The level of item difficulty is considered by further indicators that can be used to determine if the item is useful enough to be included in the academic achievement test or not. The difficulty level of an item refers to the proportion of students answering an item correctly. The item difficulty index, also known as the *p*-value, is calculated by dividing the number of students who answered the question correctly by the total number of students who answered the question (Srivastav 2000). Furthermore, the difficulty index can range from zero to one. While zero value indicates that no student answered the item correctly, one value indicates that all students answered the item correctly. The first case points out that the item is very difficult, whereas the second result explains that the item is too easy. In the present study, the item is considered difficult when the difficulty value was less than 0.25. In addition, the item is considered to be easy when the difficulty value was greater than 0.80. The results illustrated that the test items provided acceptable difficulty values ranged from 0.30 to 0.78.

On the other hand, the value item discrimination provides a suggestion of the degree to which an item correctly differentiates among the test takers on a certain domain (Whiston 2005). Steps of calculating the discrimination index were described by Kline (2005). First, those students who have the highest and lowest overall test scores are sorted into two groups: the upper group which is made up of the 25–33% who have the highest overall test scores and the lower group which is made up of the bottom 25–33% who have the lowest overall test scores. Subsequently, the *p*-value of each item for the upper and lower groups is computed.

Finally, the item discrimination index is provided simply by subtracting the *p*-values of the two groups, upper group and lower group. However, the item discrimination indices diverge from -1.00 to $+1.00$. The overall test scores were arranged in the descending method. Subsequently, the top 19 students (27% of the total number 69 of students) were included in the upper group, while the lowest 19 students (27% of the total number 69 of students) were selected in the lower group. The item discrimination indices were calculated. The items were categorized concerning their discrimination indices according to the criteria recommended by Ebel (cited by Mitra et al. 2009). The item with an index of discrimination of 0.40 and higher is considered an excellent item, while the value ranging from 0.3 to 0.39 is regarded as good item. Therefore, the item discrimination value of 0.2 to 0.29 is considered as being acceptable, whereas the ratio from 0 to 0.19 refers that test item should be revised. Moreover, the item with a negative discrimination index value should be removed from the academic achievement test. Fortunately, the findings reported that all discrimination indices for the test items were positive values.

Moreover, the results showed that the test items which provided sufficient discrimination value ranged from the good to excellent item.

Based on the results of item analysis, 61 items was the number of test questions that were found to be valid enough to evaluate student achievement in electrical engineering. Conversely, the total number of the test items was 60 items. Therefore, one item was removed from the academic achievement test.

16.4 Discussion and Conclusions

Measurement of learning performance as a means to assess academic achievement and thus the development of knowledge on an individual level is a core educational activity not only in vocational secondary schools. Only empirical data, independently of its specific format, is one of the key indicators in each educational measurement. However, measurement is often difficult, and appropriate tools need to be subject-specific and reliable. Due to that, the present study addresses the construction and validation of academic achievement tests in order to measure students learning performance in the electrical engineering. The development of the academic achievement test included a rigorous process of planning, creating, and testing. However, many of the researchers explained that careful attention must be paid to the purpose of the test, because the usefulness of test interpretations depends on the agreement with which the purpose and the domain represented by the test have been explicated (AERA et al. 1999). Therefore, the purpose, the educational content, and the target population of the academic achievement test were identified. This test aimed to assess student's learning performance in the electrical engineering for vocational secondary school in Egypt (Mabed 2013).

According to Crocker and Algina (1986), one or more actions may be applied during the development of an instrument such as content analysis, review of the literature, critical incidents, direct observations, expert judgment, and instruction objectives. The researchers had taken into account several activities to overcome the problems, which might be inherent in the test. For that reason, the test items should reflect the knowledge directly related to electrical engineering. Therefore, the content of electrical engineering was analyzed to identify the educational objectives. After experts' review has been applied to test phase and content validity, the output was a 76-item objective, reflecting five main topics in the electrical engineering. To increase the content validity of the test, the importance of each topic and the number of pages were determined during preparing the blueprint of the academic achievement test.

Those 76 items were generated to assess student's learning performance in electrical engineering. The academic achievement test items covered four taxonomy levels, which include knowledge, comprehension, application, and analysis. Presenting a clear, simple, and concise direction is an integral part of well-constructed test items (Osterlind 1998). Therefore, the first draft of the academic achievement test and the directions for the test takers were given to a panel of

experts to judge the appropriateness of them. Authors conducted the pilot in order to collect evidence that determines the validity and reliability issues. The characteristics of the pilot study sample were similar to the target group to use for the final test. The results from the pilot study suggested that the academic achievement test demonstrated acceptable reliability across the test.

The alpha coefficient for the whole test and all the sublevel was reasonable:

1. The academic achievement test showed evidence of discriminant and convergent validity.
2. The principal component factor analysis as the extraction technique and a varimax rotation as the orthogonal rotation method procedure were executed by using the statistical program PASW. The results of factor analysis revealed that the application level explained the higher proportion of the variance, while the comprehension level accounted for the lower percent of the variance.
3. Moreover, the academic achievement test had a good item difficulty falling within the range of 0.30 to 0.78 with a discriminative index of the test items ranging from 0.41 to 0.79. Since the academic achievement test items were designed to reflect predetermined objectives, this variation in value to the items would be expected.

In sum, the findings indicated that the 60-item test holds promise as a valid and reliable academic achievement test to measure student's learning performance in electrical engineering. With a duration of approximately 100 min for answering the whole test, it may be applied in the context of regular educational activities in vocational secondary schools. The calculation of the scores is supported by an easy to be used template.

Finally yet importantly, the construction and development of the academic achievement test in the present study should be considered in the light of a few limitations. The academic achievement test is related to electrical engineering content in the vocational education school. Moreover, although the research sample of pilot study was appropriate for the correlation matrix and factor analysis, it was relatively small. Future research utilizing a larger sample size is necessary to provide evidence regarding the test validity.

References

- Adams, J. (2014). The language of assessment. In M. McDonald (Ed.), *The nurse educator's guide to assessing learning outcomes* (3rd ed., pp. 11–30). Burlington: Jones & Bartlett Learning.
- Algarabel, S., & Dasí, C. (2001). The definition of achievement and the construction of tests for its measurement: A review of the main trends. *Psicológica*, 22, 43–66.
- American Educational Research Association AERA, American Psychological Association APA, & National Council on Measurement in Education NCME. (1999). *The standards for educational and psychological testing*. Washington, DC: AERA.
- Ariyo, A. O. (2007). Construction and validation of a general science aptitude test (GSAT) for Nigerian junior secondary school graduates. *Ilorin Journal of Education*, 27, 20–29.

- Bayrak, B. K., & Bayram, H. (2010). The effect of computer aided teaching method on the students' academic achievement in the science and technology course. *Procedia – Social and Behavioral Sciences*, 9, 235–238. <https://doi.org/10.1016/j.sbspro.2010.12.142>.
- Bloom, B. S., Engelhard, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R.. (1956). *Taxonomy of educational objectives: The classification of educational goals* (Handbook I: Cognitive Domain). New York: David McKay.
- Carle, A. C., Jaffee, D., & Miller, D. (2009). Engaging college science students and changing academic achievement with technology: A quasi-experimental preliminary investigation. *Computers & Education*, (52), 376–380. Retrieved from <http://www.elsevier.com/locate/compedu>
- Choy, J. L. F., O'Grady, G., & Rotgans, J. I. (2012). Is the Study Process Questionnaire (SPQ) a good predictor of academic achievement? Examining the mediating role of achievement-related classroom behaviours. *Instructional Science*, 40(1), 159–172. <https://doi.org/10.1007/s11251-011-9171-8>.
- Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3–25). Mahwah: Lawrence Erlbaum Associates Publishers.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>.
- Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-graph: Tutorial and annotated example. *Communications of the Association for Information Systems*, 16(1), 91–209.
- Haislett, J., & Hafer, A. A. (1990). Predicting success of engineering students during the freshman year. *The Career Development Quarterly*, 39(1), 86–95. <https://doi.org/10.1002/j.2161-0045.1990.tb00238.x>.
- Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal*, 20(2), 195–204. [https://doi.org/10.1002/\(SICI\)1097-0266\(199902\)20:2<195::AID-SMJ13>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1097-0266(199902)20:2<195::AID-SMJ13>3.0.CO;2-7).
- Kline, T. J. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks: SAGE Publications.
- Köhler, T., & Mabed, M. (2014). Academic achievement in vocational secondary schools. Construction and validation of a test to assess potential learning performance. In G. A. Papadopoulos (Ed.), *Proceedings of the 9th international Conference on Knowledge, Information, and Creativity Support Systems (KICSS)*. Nicosia: Cyprus Library.
- Mabed, M. (2013). *The effectiveness of using blended learning in teaching electrical engineering for industrial education students*. Dissertation at TU Dresden, Dresden.
- Mabed, M., & Köhler, T. (2016). Aligning performance assessments with standards: A practical framework for improving student achievement in vocational education. In S. Kunifuji, G. A. Papadopoulos, A. M. J. Skulimowski, & J. Kacprzyk (Eds.), *Knowledge, information and creativity support systems* (Volume 416 of the series Advances in intelligent systems and computing). New York: Springer.
- Mitra, N. K., Haleagrahara, N., Ponnudurai, G., & Judson, J. P. (2009). The levels of difficulty and discrimination indices in type A multiple choice questions of pre-clinical semester 1 multidisciplinary summative tests. *International E-Journal of Science, Medicine & Education*, 3(1), 2–7.
- Osterlind, S. J. (1998). *Constructing test items: Multiple-choice, constructed-response, performance, and other formats* (Evaluation in education and human services, 2nd ed.). Boston: Kluwer Academic Publishers. Retrieved from <http://www.springer.com/la/book/9780792380771>

- Quellmalz, E., & Hoskyn, J. (1995). Classroom assessment of reasoning strategies. In G. D. T1-Contributors A2-Phye (Ed.), *Handbook of classroom assessment: Educational psychology* (pp. 103–130). San Diego: Academic. <https://doi.org/10.1016/B978-012554155-8/50007-7>.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417–458. <https://doi.org/10.1111/j.1468-0262.2005.00584.x>.
- Romano, J., Wallace, T. L., Helmick, I. J., Carey, L. M., & Adkins, L. (2005). Study procrastination, achievement, and academic motivation in web-based and blended distance learning. *The Internet and Higher Education*, 8(4), 299–305.
- Selçuk, G. S., Sahin, M., & Açıköz, K. Ü. (2011). The effects of learning strategy instruction on achievement, attitude, and achievement motivation in a physics course. *Research in Science Education*, 41(1), 39–62. <https://doi.org/10.1007/s11165-009-9145-x>.
- Shukla, R. (2005). *Dictionary of education*. New Delhi: A.P.H. Pub.
- Smith, D. R. (2001). Wechsler individual achievement test. In J. J. Andrews, D. H. Saklofske, & H. L. Janzen (Eds.), *Handbook of psychoeducational assessment: Educational psychology* (pp. 169–193). San Diego: Academic. <https://doi.org/10.1016/B978-012058570-0/50008-2>.
- Srivastav, P. G. N. (2000). *Management of teacher education: A handbook*. New Delhi: Concept Publishing.
- Whiston, S. C. (2005). *Principles and applications of assessment in counseling* (2nd ed.). Belmont: Thompson, Brooks Cole.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

