We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Developing and Evaluating Educational Programs

B. Charles Tatum

Abstract

The chapter describes a system for the development and evaluation of educational programs (e.g., individual courses or whole programs). The system describes steps that reflect best practices. The early stages in development (planning, design, development, implementation) are described briefly. The final stage (evaluation) is described in more detail. The evaluation step is a four-tiered process based on the Kirkpatrick and Kirkpatrick model and a performance indexing measurement system of Tatum and Nebeker. The chapter should be a valuable guide for teachers, program directors, and department chairs in their efforts to create and maintain quality educational experiences and high levels of student learning.

Keywords: educational programs, curriculum development, course evaluation, program evaluation, student learning

1. Introduction

Most educational endeavors (e.g., producing curricula, programs, courses) follow a pretty standard set of activities for the purpose of educating students as shown in **Figure 1**. The chapter will rely mostly on college and university curriculum examples, but this does not exclude primary and secondary schools. Think of this as a roadmap. Like any roadmap, it is not the only way to get from Point A to Point B, but it will show the landscape and road signs from which to navigate through the process of creating new and better educational experiences for students. This chapter will briefly describe the first four phases of the process, and then focus in more detail on the evaluation phase. The emphasis on the evaluation phase is in line with current trends in education that view student learning and success as essential to academic performance.

Figure 1.
Sequence of activities in educational planning, designing, developing, implementing and evaluating programs and courses.

2. Phase I: planning

Planning is the first of five phases in creating an educational experience (e.g., an individual course, an academic program). Planning includes a set of data gathering

and assessment activities aimed at helping to decide whether or not to proceed to the design, development, implementation, and evaluation phases. The output of the planning phase should be a written concept proposal that makes an academic case for proceeding to the subsequent phases. This planning document can then be submitted to the appropriate approval structures (e.g., principal, department chair, dean, academic committee).

A planning document needs to cover several areas, including but not limited to, a mission statement, a needs analysis, required resources, benchmark assessments, general target competencies and/or outcomes, and an evaluation plan. Before proceeding to the design phase, a few words should be said about the distinction between competencies and outcomes and the evaluation plan.

2.1 Competencies

Competencies refer to a general set of knowledge, skills, abilities, and other personal traits (e.g., attitudes, ethics, interests) that predict behavior in a wide variety of situations. Competencies provide the student with an integrated "mental model" of the current state and evolving standards of the field: [1–8]. Examples of competencies include problem solving ability, communication skills, personal and professional ethics, and values, to name just a few.

2.2 Outcomes

Outcomes come in two varieties: program learning outcomes (PLOs) and course learning outcomes (CLOs). Learning outcomes tend to be more specific than competencies, with PLOs representing broad program objectives and CLOs representing specific ways in which a particular course meets those objectives [9]. Learning outcomes should be expressed as observable, behavioral outcomes (i.e., what the student is expected to do), and typically include an action verb and a target content area. The action verb is often taken from Bloom's taxonomy [10, 11] that ranges from low level actions (e.g., remember, understand) to high level actions (produce, construct). For example, a PLO might be: A graduate of this program will be able to evaluate research designs and construct research projects. A CLO for a research course in the program might be: At the completion of this class, the student will be able to identify the major designs from Campbell and Stanley [12]. Another CLO for the same class might be: At the completion of this class, the student will be able to create a research project using one of the Campbell and Stanley designs.

2.3 Evaluation plan

An important element in any planning document is an answer to the question: What will be used as evidence that a program or course was successful? One highly researched and successful approach addressing this question comes from the model proposed by Kirkpatrick and Kirkpatrick [13]. The model identifies four evaluation levels as shown in **Table 1**. These levels are (a) reaction: participant satisfaction and self-assessment of learning, (b) learning: the learners' knowledge and skill improvement, (c) behavior: transfer of learned skills to other areas (e.g., jobs or future classes), and (d) results: impact on the institutions success and improvement. Often, educators seem satisfied with only assessing the first two of these levels (reaction and learning). The last two (behavior and results), however, may be even more essential to academic performance. The last two levels go beyond just learning, and assess what students can do and how this contributes to a more general measure of educational success.

Evaluation Level	Examples of Evaluation Source and Data Type
1: Reaction	• Participant ratings of satisfaction with materials, activities, instructors, classes, facilities, and general program administration • Student self-assessment of learning
2: Learning	• Exam scores • Course grades • Demonstrated skill level of participants from skill-based assessments (e.g., observations, expert ratings, simulations)
3: Behavior	• Student self-report of skills used on a job • Supervisor or manager ratings of learned job skills • Ability to apply learned knowledge or skill in subsequent courses
4: Results	• Number of program applicants • Number of successful participants • Positive reviews from school districts or external agencies • Increased number of graduates gaining employment or advancing to other programs

Table 1.
Sample outline of an evaluation plan with Kirkpatricks' four levels.

3. Phase II: design

The design phase involves creating a general structure for later development (see [14, 15]). Completing these steps will help guide the next phase (Development). Several actions should be taken such as (a) establish time frames for the future phases, (b) specify desired competencies or learning outcomes, (c) identify learning and performance activities that demonstrate successful achievement of the competencies/outcomes, (d) set prerequisites (e.g., students taking Algebra II must have completed Algebra I, students enrolled in a college program must have a high school diploma), (e) determine the major administrative concerns, and (f) decide what data can be collected that will reflect the four Kirkpatrick levels (see **Table 1** for some suggestions).

For some ideas about what design actions can be taken, see the checklist below (adapted from [14]).

- Who is the primary point of contact (POC)?

- To whom are the applications submitted?

- How will candidates and participants be kept informed?

- How will prerequisites be assessed?

- Who will ensure the application materials are complete?

- Who reviews and approves the applications?

- When and where will training be conducted?

- Where will the student records (e.g., attendance, course completion, start dates) be kept?

- What budget will pay for the support personnel?

- How will exams be administered?

- How will exams be secured?

- Who will write, proctor, and grade the exams?

- Where, when, and how will skills training be conducted?

- What corrective action steps will be used and who will monitor this process?

- What awards and or recognitions will be issued?

- Who will oversee ongoing program maintenance?

- What sources of data are required to assess the Kirkpatrick four levels and how will they be obtained?

4. Phase III: development

The Development phase described below explains the steps required to produce an educational practice that is ready for implementation (Phase IV). The development phase can generally be carried out in two steps.

4.1 Select and develop learning and performance activities

This step builds on the work completed earlier under Design (identifying learning and performance activities). This is where the actual learning and training activities are generated and matched to the learning outcomes/competencies. There are two, not mutually exclusive, options for achieving this step: (a) find relevant learning and performance activities from external sources, and (b) select or develop these activities in-house.

Procuring the relevant activities from an external source is far less time consuming than developing them in-house. The principal disadvantage is that the learning opportunities offered by outside sources may not be entirely suitable for the curriculum (i.e., the activities may not address the outcomes and competencies in the most direct and relevant fashion).

Selecting and developing the learning and performance activities in-house allows for customized experiences that can target specific knowledge, skills, and abilities. Home-grown educational experiences have the advantage of being directly relevant to the outcomes and competencies identified for the curriculum. The disadvantage of this customized approach is that it can be very specialized and may require a high degree of instructional design expertise and technical skill to develop.

4.2 Establish tests and measures of outcomes/competencies

In this step, the learning outcomes/competencies that the students are acquiring will be assessed. The process of developing tests and measures is described in the literature on testing theory and practice (see [14, 16]). The easiest approach is to locate existing tests and measures. These existing materials may come from a variety of sources including curricula from other institutions, training classes, certification programs, continuing education units, extension classes, and so forth. Below is a partial list of candidate tests and measures that can be used to assess whether students are mastering the content and meeting the outcomes and/or competencies:

- Participant satisfaction ratings

- Participant self-assessment of learning

- Course grades

- Class quiz scores

- Midterm and final exam scores

- Instructor ratings of class assignments

- Final project/thesis evaluation

- Supervisor's assessment ratings

- Peer evaluations

- Self-review of functional skills

- Expert ratings of oral presentations

- Simulation/game scores

- Skill exercise observations

- Panel review recommendations

- Portfolio analyses

When an adequate set of existing assessment tools cannot be located from external sources, then customize tests and measures must be developed. When developing custom-assessment items, two important criteria must be met—reliability and validity [16]. A reliable assessment is one that is consistent. A valid assessment is one that is accurate. The first criterion (reliability) is generally established by showing that the test or measure is stable over time (e.g., repeated use of class quizzes yield consistent scores). The second criterion (validity) assures that the tests or measures accurately evaluate what they are intended to appraise. There are several techniques for ensuring the validity of tests and measures, but the most common validity check is to use Subject Matter Experts (SMEs) who closely examine the tests and measures and form a consensus that these tools in fact reflect the relevant outcomes or competencies.

5. Phase IV: implementation

After a course or program has been developed, it is ready to be implemented. There is no standardized process for implementation, but educational institutions have developed and implemented initiatives across a wide variety of disciplines and there is a large body of common practices to draw from: [14, 15, 17, 18]. In general, there are at least four steps involved in a standard implementation.

5.1 Conduct pilot studies

A pilot study is a "pre-study" conducted as a dry-run prior to launching the full effort. The study should be on a much smaller scale than the full curriculum (e.g., fewer students, less costly technology, fewer classes), but still preserve the essence of the program.

5.2 Refine essential elements

The results of the pilot study should be examined and lessons learned should be noted. Specifically, at least the following elements need to be reviewed and modifications made.

5.2.1 Outcomes/competencies

Are these the right outcomes or competencies for this curriculum? Should more be added? Should some be deleted?

5.2.2 Time frames

Is the timing of events (course duration, project times, testing schedules) optimal? Where can changes occur?

5.2.3 Prerequisites

Were the correct prerequisites identified? Should some be added? Should some be removed?

5.2.4 Administrative procedures

Was the administration of the pilot study efficient? Where were the administrative bottlenecks and glitches? How can these be improved?

5.2.5 Learning and performance activities

Did the learning and performance activities produce the intended outcomes? Should new activities be added? Should some activities be discarded? Can improvements be made to the existing set of activities?

5.2.6 Tests and measures

Did the knowledge tests and performance measures assess the outcomes and competencies of the students as expected? What adjustments should be made?

5.2.7 Data collection

Are the data collected easily obtained and in a usable form? Can clear conclusions be drawn from these data?

5.3 Market the initiative

To help ensure success, a marketing plan should be devised to advertise the program and recruit students. The following items should be considered: (a) identify

the target audience, (b) align the marketing objectives with the curriculum objectives, (c) create a communication plan, (d) publish a schedule, and (e) use specific institutional marketing techniques (e.g., fact sheets, web and electronic media, newsletter, brochures, communication networks, open house events, personal visits to potential recruiting venues).

5.4 Launch full curriculum

In this step, the program gets implemented in accordance with the pilot study modifications.

6. Phase V: evaluation

After the course or program has been implemented, it must be evaluated for effectiveness. This evaluation should be driven by some formal model such as the Kirkpatrick and Kirkpatrick model [13] shown in **Table 1**. If the Kirkpatrick's model is adopted, then data are required that assess each of the four levels. If the evaluation is for a single course, then the tests and measures will be mainly, but not exclusively, relevant to level 2 (learning). If the evaluation is for an entire program, then all the levels should be assessed (as shown in **Figure 3** to be discussed below). I'll begin with the evaluation of a single course, and show one possible approach.

6.1 Evaluating a course

Once the knowledge tests and performance measures have been administered to students in a class, each person should have a set of relatively objective scores. These scores, when combined, should show how successful the student was with regard to the class outcomes. The approach illustrated here is based on "performance indexing" developed by Tatum and Nebeker [19]. Performance indexing is a system for combining and weighting a set of scores and generating an overall index. Performance indexing has been employed successfully in fields outside of education (e.g., real estate, environmental quality management, organizational improvement), but can be used just as effectively in an educational setting. The weighting feature is especially useful because it takes into account how valuable each test or measure is in the overall assessment. If, for example, in a biology class, mid-term and final exam performance is more important than homework, this difference will be reflected in the final index. Often, the degree of importance is reflected by the number of points that can be earned by each assignment. Performance indexing offers a more sophisticated system for balancing performance and getting at the essence of student learning. An example of performance indexing used in a hypothetical class is shown in **Figure 2**.

There are several steps to developing and using performance indexing (for a more complete discussion of the topic see Tatum and Nebeker [19]). The most essential features are (a) each test or measure is given a weight according to its importance in the assessment, and (b) an overall index score is generated from the weighted values (e.g., 370–400 is outstanding performance). **Table 2** is a step-by-step guide for building and using the performance index table in **Figure 2**.

6.2 Evaluating a program

A program (e.g. clinical psychology, biology, history) is designed so that students graduate having met certain competencies or PLOs. Whether the program uses

Figure 2.
A performance index table used to assign a final course grade to a single student in a hypothetical class.

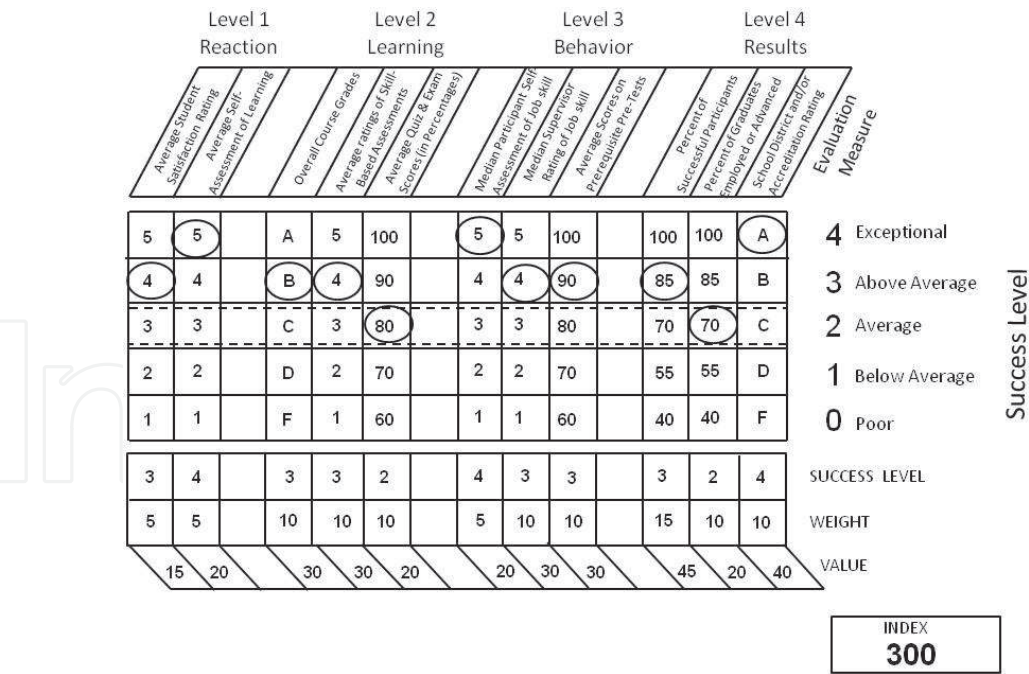| Steps	Building a Performance Index Table
One	Record the Course Learning Outcomes (CLOs) at the top of the table.
Two	List the specific tests and measures for each CLO in the diagonal spaces.
Three	Identify an average (satisfactory) score for each measured course requirement (center row labeled "satisfactory").
Four	Identify two score levels above the satisfactory level (e.g., the highest possible score is "outstanding," and midway from satisfactory to the highest is "exceptional").
Five	Identify two score levels below satisfactory (e.g., lowest possible score is "failed," and midway between lowest and satisfactory is "marginal").
Six	Use a 100-point allocation system to weight each requirement. Place these weights in the second to the last row ("weight"). The weights should add to 100 with the larger numbers representing higher importance.
Steps (continued)	Using the Performance Index Table
Seven	Collect data from each student and determine the success level. Each circled number in the table represents the score the student earned for that specific requirement. The score corresponds to a "success level" to the right (0-4). The level (0-4) is entered on the third row from the bottom also labeled "success level."
Eight	Multiply the success level (third row from the bottom) by the weight (second row from the bottom) to obtain a value for each requirement (last row).
Nine	Add all the Values to obtain an Index. For example: • An Index between 370 and 400 signifies outstanding performance (A) • An Index below 93 signifies failure (F).* *The grade ranges are suggestions only. Each program or instructor should select the ranges best suited for them.

Table 2.
Steps for building and using a performance index table.

competencies or PLOs is a matter of preference, but regardless of this choice, the CLOs must be designed to meet one or more of these competencies or PLOs. When students successfully complete all of courses in the program, they will have satisfied the expected objectives of the program and will leave with specific knowledge, skills, abilities and other desired characteristics (e.g., attitudes, personal ethics, interests).

Figure 3 is an example of how performance indexing can be used to evaluate an entire program (as opposed to a specific course within that program as shown

Figure 3.
A performance index table used to evaluate a hypothetical academic program.

above in **Figure 2**). Developing an index table for a program involves basically the same steps outline in **Table 2**, with a few modifications. The evaluation measures (shown in the diagonal spaces) are based on the Kirkpatrick and Kirkpatrick [13] levels (see **Table 1**), which are shown at the top of the **Figure 3**. The specific evaluation measures will vary from program to program, but each measure should fall into one of the levels. For example, level 1 (reaction) is supposed to indicate the students' satisfaction rating of the program and an assessment of how much they think they have learned. These ratings can be obtained from each class or as part of an exit survey at the end of the program. Level 2 (learning) is intended to reveal, on a more objective basis, how much the students learned (in this case based on grades, ratings of acquired skills, and test scores average across classes). Level 2 is closely tied to the competencies or PLOs of the program. Level 3 (behavior) is an indication of the degree to which the program changed the student's behavior and the extent to which the student can transfer this behavior to other settings (e.g., did they learn valuable job skills, did the acquire knowledge and skills in prerequisite classes that they can apply to future classes?). Although level 3 is normally associated with assessing an entire program, it is still possible to include behavioral measures at the course level. For example, **Figure 2** shows an "internship supervisor rating" as an essential measure of the student's ability to apply what was learned. Finally, level 4 (results) is supposed to show that the program had a positive impact on the current success and future improvement of the institution (e.g., local school, school district, college). Evidence for positive results can be demonstrated by a variety of data such as graduation rates, employment success, advancement to higher levels of education, or ratings by external agencies. Level 4 measures are not common among educational institutions when evaluating individual programs (although these data are routinely collected at higher levels), but they should be. To capture the essence of academic performance, we must assess the degree to which our programs contribute to the general success and welfare of the broader academic community.

Once the final index is computed, the overall success of the program can be evaluated. In the hypothetical program depicted in **Figure 3**, the index score of 300 indicates that the program is above average. A close examination of **Figure 3** will also

reveal where the program is performing well (students rate their learning and job skills as exceptional, the program gets an exceptionally high rating by external agencies) and where it requires improvement (test scores are low, there is a low percentage of students finding jobs or advancing to other programs).

7. Concluding remarks

The phases and steps advocated here are obviously a mechanistic (non-theoretical) approach. It resembles Tyler's [20] thinking about curricular design more than contemporary thought (e.g., [21, 22]). There is nothing wrong with a more mechanical approach. In fact, the phases and steps proposed in this article are not incompatible with modern views of education such as the sharing of common goals [23], scaffolding [24], or the spiral curriculum [25]. At some point, however, we need to find and follow a path towards building an educational program, and this roadmap shows us the way without too many detours.

Academic performance has been the focus of much research and interest during the past few years. Initiatives such as No Child Left Behind [26] and Race to the Top [27] have generated much debate and concern regarding the components of academic performance [28] and the optimal methods for assessing learning and success [29]. This chapter proposes a method for developing and evaluating courses and programs that gets at the heart of academic performance in five phases (i.e., planning, design, development, implementation, and evaluation). The first four phases are a prelude to what the author considers the true essence of academic performance; namely, the identification and measurement of performance indicators. This chapter presents an evaluation model (based on [13, 19]) that guides the user down a well-traveled road that leads, in the end, to a quantitative understanding of student course performance and program success. In the inimitable words of Peter Drucker: *You can't manage what you can't measure*.

Acknowledgements

Author details

B. Charles Tatum
Department of Psychology, College of Letters and Sciences, National University, La Jolla, CA, USA

*Address all correspondence to: ctatum@nu.edu

IntechOpen

References

[1] Cummings T, Worley C. Organizational Development and Change. Cincinnati, OH: Southwestern College; 2001

[2] Hooper L, Begg M, Sullivan L. Integrating competencies and learning outcomes in core courses for the MPH. Public Health Reports. 2014;**129**:376-381

[3] Kuh G, Ewell P. The state of learning outcomes assessment in the United States. Higher Education Management and Policy. 2010;**22**(1):1-20

[4] McClelland D. Testing for competence rather than "intelligence". The American Psychologist. 1973;**28**:1-14

[5] Office of Personnel Management. Assessment and evaluation: Competency modeling. 2015. Available from: http://www.opm.gov/services-for-agencies/assessment-evaluation/competency-modeling/

[6] Spencer L, Spencer S. Competence at Work: A Model for Superior Performance. New York, NY: Wiley; 1993

[7] Worley C, Rothwell W, Sullivan R. Competencies of OD practitioners. In: Rothwell W, Sullivan R, McLean G, editors. Practicing Organization Development: A Guide for Consultants. 2nd ed. San Diego, CA: Pfeiffer; 2005. pp. 135-163

[8] Worley C, Varney G. A Search for a Common Body of Knowledge for Master's Level Organization Development and Change Programs. Winter ed. Academy of Management ODC Newsletter; 1998. pp. 1-4

[9] Center for Teaching Support and Innovation. What are learning outcomes? 2019. Available from: https://teaching.utoronto.ca/teaching-support/course-design/developing-learning-outcomes/what-are-learning-outcomes/

[10] Anderson L, Krathwohl D. Bloom B. a Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York, NY: Longman; 2001

[11] Bloom B, Engelhart M, Furst E, Hill W, Krathwohl D. Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain. New York, NY: David McKay; 1956

[12] Campbell D, Stanley J. Experimental and Quasi-Experimental Designs for Research. Chicago, IL: Rand McNally; 1963

[13] Kirkpatrick D, Kirkpatrick J. Implementing the Four Levels: A Practical Guide for Effective Evaluation of Training Programs. San Francisco, CA: Berrett-Koehler; 2007

[14] Hale J. Performance Based Certification: How to Design a Valid, Defensible, Cost-Effective Program. San Francisco, CA: Jossey Bass; 1999

[15] White A. Building an Internal Certification Program. Alexandria, VA: ASTD Press; 2006. Available from: http://books.google.com/books?id=M99L0vHQ5MkC&dq=Building+an+Internal+Certification+Program&printsec=frontcover&source=bl&ots=mmfkhdCJNn&sig=INnJhWrUba5A263EyjbAAHyOlys&hl=en&ei=ZePDSafLBJjmnQfoprlE&sa=X&oi=book_result&resnum=3&ct=result#PPP1,M1

[16] McDonald R. Test Theory: A Unified Treatment. Mahwah, NJ: Lawrence Erlbaum; 1999

[17] Kotler P. Marketing Management: Analysis, Planning, Implementation,

and Control. Englewood Cliffs, NJ: Prentice Hall; 1988

[18] Rossi P, Freeman H, Lipsey M. Evaluation: A Systematic Approach. Thousand Oaks, CA: Sage; 1999

[19] Tatum B, Nebeker D. A strategic approach to measuring organizational performance. In: Loman R, editor. Handbook of Organizational Consulting Psychology. San Francisco, CA: Jossey-Bass; 2002. pp. 692-729

[20] Tyler R. Basic Principles of Curriculum and Instruction. Chicago, IL: University of Chicago Press; 1949

[21] Knight P. Complexity and curriculum: A process approach to curriculum-making. Teaching in Higher Education. 2001;**6**(3):369-381

[22] Parker J. Reconceptualising the curriculum: From commodification to transformation. Teaching in Higher Education. 2003;**8**(4):529-543

[23] Stenhouse L. An Introduction to Curriculum Research and Development. London, UK: Heineman; 1975

[24] Wildman T. Taking Seriously the Intellectual Growth of Students: Accommodations for Self-Authorship. New Directions for Teaching and Learning, no. 109. Wiley Periodicals; 2007. DOI: 10.1002/tl.262. Available from: sss.interscience.wiley.com

[25] Bruner J. The Process of Education. Cambridge, MA: Harvard University Press; 1996

[26] U.S. Department of Education. The No Child Left Behind Act of 2001. Available from: https://www2.ed.gov/nclb/overview/intro/execsumm.pdf

[27] U.S. Department of Education. Race to the Top. 2009. Available from: https://www2.ed.gov/programs/racetothetop/factsheet.html

[28] Tatum B A predictive model of academic performance. In Haumann R, Zimmer G, editors. Handbook of Academic Performance: Predictors, Learning Strategies and Influences of Gender. Hauppauge, NY: Nova Science; 2013. pp. 109-126

[29] York T, Gibson C, Rankin S. Defining and measuring academic success. Practical Assessment, Research and Evaluation. 2015;**20**(5):1-20