A Comparison and Classification of Grading Approaches used in Engineering Education

A. R. Carberry[1]
Arizona State University
Mesa, Arizona, USA
https://orcid.org/0000-0003-0041-7060

S. A. Atwood
Elizabethtown College
Elizabethtown, Pennsylvania, USA
https://orcid.org/0000-0001-9759-9328

M. T. Siniawski
Loyola Marymount University
Los Angeles, CA, USA
https://orcid.org/0000-0001-5999-9988

H. A. Diefes-Dux
University of Nebraska
Lincoln, NE, USA
https://orcid.org/0000-0003-3635-1825

Conference Key Areas: Talent Management and Strong Demand for Democratic Involvement in Educational Processes

Keywords: grading systems, assessment, engineering higher education

[1] A. R. Carberry, adam.carberry@asu.edu

ABSTRACT

Grades are intended to communicate achievement associated with a learning experience. Engineering educators in higher education often default to a particular grading approach without considering how the approach impacts student achievement. This work proposes a model for comparing and classifying commonly used grading systems in engineering higher education. Examples from the engineering education literature revealed five general categories of grading: 1) normative, score-based grading, 2) summative grading, 3) standards-based grading, 4) mastery-based grading, and 5) adaptive grading. (Note: variations in naming conventions were observed.)

Each grading system was examined to determine key characteristics of the system and how student performance was ultimately assessed. A continuum of grading approaches was created after discovering that each system ranged in its intention to select and/or develop talent. The most widely adopted approaches to grading in engineering higher education, norm-based grading, were classified using purely selective processes (e.g., letter grades). Alternative, learning outcomes-based grading approaches differentiate themselves by the level in which they attempt to develop talent. This was determined by examining differences in how the grading system impacted sequencing of content, course pace, number of attempts to demonstrate achievement, scale and weight of performance, feedback provided, and basis for a final grade. The resulting continuum provides a tool for engineering educators to compare and discuss grading approaches in order to select an appropriate system for their course or program. Informed decisions on grading can have a critical impact in student retention and program improvement.

1 INTRODUCTION

Assessment of student achievement using a grading system is a major task asked of educators since the late 1700s (Postman, 1992). Engineering instructors are faced with the constant challenge of selecting appropriate assessment measures and grading systems for the courses they teach. Take for example Dr. Smith.

> *Dr. Smith is an engineering instructor with a clear idea of and confidence in her ability to teach the course's topics. Her approach to assessing her students is the same way she used to be assessed as a student herself. This involves summing scores on a variety of assignments - homework, quizzes, projects, and exams - and calculating a final grade. The process, grading scale, and learning outcomes associated with the course are outlined on her syllabus. Dr. Smith's reflection on her grade distributions at the end of the course cause her to question whether the assignments and grading system accurately represent how well the students have learned the course material. She is interested in making some changes, but is unaware of who to ask or where to go to learn about alternatives beyond what is familiar.*

This paper provides engineering instructors like Dr. Smith and engineering programs with the information needed to make an informed decision about how student learning will be assessed in their classrooms.

Grades are intended to communicate how well students have achieved the learning outcomes established for a course of study by: 1) communicating achievement to other interested parties, 2) selecting, identifying, or grouping students for certain educational pathways or programs, 3) providing information to students for self-evaluation/assessment, 4) providing incentives for students to learn, and 5) documenting student performance for the purpose of evaluating the effectiveness of instructional programs (Guskey & Pollio, 2012; Muñoz & Guskey, 2015). The wide variation in motivations for and uses of grades leads to confusion over their primary purpose in higher education. Should grades be used for ***selecting*** talent by differentiating student performance for outside parties or internal programs or for ***developing*** talent by improving performance through instructor feedback and student self-assessment?

This paper compares and classifies the variety of grading approaches used in engineering education to provide a continuum for engineering educators and programs to differentiate systems and make an informed decision on what approach to use in their course(s).

2 BROAD CATEGORIES OF GRADING SYSTEMS

2.1 Norm-based Grading

The most widely used systems in higher education are normative, score-based approaches (norm-based) that aim to differentiate and ***select*** talent (Freeman & Lewis, 1998; Walvoord & Anderson, 1998; Huba & Freed, 2000; Morgan, Dunn, Parry, & O'Reilly, 2004; Muñoz & Guskey, 2015; Stevens & Levi, 2004; Suskie, 2004). Such grading systems efficiently maximize differences between student achievement by measuring student performance on a number of discrete and disparate tasks throughout the semester. This approach makes it difficult for students to clearly understand what constitutes success. Conflicting evidence about student achievement and progress is often lost in the process of weighting or combining grades on separate assignments (Cross & Frary, 1999). The resulting end-of-semester grade ultimately communicates a student's performance on separate tasks relative to other students rather than individual achievement referenced directly to the stated course learning outcomes (Muñoz & Guskey, 2015, Angelo & Cross, 1993; Broad, 2000; Sadler, 1987, 2005 & 2009; Shay, 2005). The learning outcomes referred to are defined as explicit statements of what students should be able to do if they have learned what their instructor has attempted to teach them (Diamond, 2008; Felder & Brent, 2016). This *selection* approach fosters negative competition among students and a belief by instructors that not all students *should* achieve mastery (Krumboltz & Yeh, 1996).

Little variation exists among norm-based grading approaches. All systems use a summation of scores or summative approach. Some include a norming process that allows for direct comparison and "curving" of grades based on the class performance as a whole.

2.2 Learning Outcomes-based Grading

Outcomes-based, objectives-based, or criterion-referenced approaches (heretofore referred to as learning outcomes-based) are alternatives that shift the goal of assessment more toward student talent development. This is accomplished through direct measures of student proficiency on well-defined course learning outcomes (Heywood, 2014; McIntyre-Hite, 2016; Sadler, 2005). The focus is on the detailed descriptions of what a student must be able to do at the conclusion of a course, which is different from course goals - broad definitions of student competence - and course learning objectives - content an instructor will cover in a course (Diamond, 2008).

These learning outcomes-based (LO-based) approaches, developed in the 1970's (Burke, 1989; Heywood, 2016; Spady, 1977), allow instructors to support students' mastery of clearly articulated course learning outcomes by aligning course activities and assessments (Heywood, 2014; Sadler, 2005). The outcomes are supplemented by specific and targeted feedback. Feedback is a critical mechanism used to increase transparency of the criteria for success and to guide students in their efforts to succeed (Butler, 1988; Guskey, 1997; Wiggins & Tighe, 1998; Post, 2014). Multiple assessment opportunities of learning outcomes offer students the ability to improve their learning. Alignment of assessment with the identified course learning outcomes increases the overall validity, reliability, transparency, and fairness of the grading process (Muñoz & Guskey, 2015; Wiggins & McTighe, 1998).

Numerous variations of LO-based grading have been found in the literature and are used in practice, including standards-based, mastery-based, competency-based, and adaptive grading. Each of these are distinguished by implementation features including point scale, pace, and opportunities to demonstrate mastery. The variation in LO-based approaches seen in the literature demonstrates an instructor's ability to select the structure that best fits their course content, teaching style, and students.

3 THE NEED FOR A TOOL TO CLASSIFY GRADING SYSTEMS

Norm-based grading systems are still highly used in engineering higher education, but the use of LO-based grading systems is on the rise as demonstrated by many examples discussed in the recent engineering education literature. For example, Henri et al. (2017) conducted a review of competency-based learning tools assessments and recommendations. Additionally, the authors of this paper have extensively examined the use and implementation of standards-based grading in various contexts (Atwood et al., 2014; Carberry et al, 2012; Sinawski et al., 2012; Hylton & Diefes-Dux, 2016) to reveal best practices (Lee et al., 2018) and student use of feedback (Diefes-Dux, 2018).

There is little confusion as to what classifies as norm-based grading. The same cannot be said for LO-based grading efforts, which have been quite scattered and bogged down by inconsistencies in nomenclature. For example, Post (2014 & 2017) highlights the use of what he refers to as standards-based grading in fluid mechanics (2014) and thermodynamics (2017) courses. The use of standards-based grading in these examples highlights an approach that leverages a two-point, pass/fail scale for nine and eleven course objectives, respectively. This binary approach to grading and multiple opportunities to achieve a score of 'pass' is inconsistent with our use of standards-based grading.

Bekki, Dalrymple & Butler (2012) also use a two-point, pass/fail scale as described in their implementation of mastery-based grading in an undergraduate engineering curriculum. Students are provided with two chances to demonstrate mastery of well-defined learning objectives, which roughly aligns with Post's description of standards-based grading. Another example labelled as mastery or competency-based grading is provided by DeGoede (2018) within an engineering dynamics course. Students must demonstrate mastery on eleven skills, but are assessed using a 5-point scale. Each skill is assessed individually using a test question and students are provided with multiple opportunities to display such mastery. This approach appears to overlap or mix with what was previously described as standards-based and mastery-based.

These examples do not introduce forms of computer assistance, but instructors are beginning to add such technology either in class or online (e.g.,Padhye & Blumenstein, 2017). This added level of complexity further muddies the water in terms of being able to compare and classify these emerging grading systems in engineering education. The continual emergence of learning outcomes-based grading system variations suggests a clear need to examine what has/is being done to help engineering instructors and programs navigate the growing grading landscape.

4 USING PRINCIPLES TO DIFFERENTIATE GRADING SYSTEMS

All grading systems are designed to assess student performance within a learning environment. A set of principles, suggested by Heywood (2016), can be used as a first level of filtration to differentiating grading systems used in engineering education. Fig. 1 presents these principles mapped directly to norm and learning outcomes-based systems introduced in the previous sections. A clear dividing line can be made separating principles and grading systems that focus on selecting and developing talent.

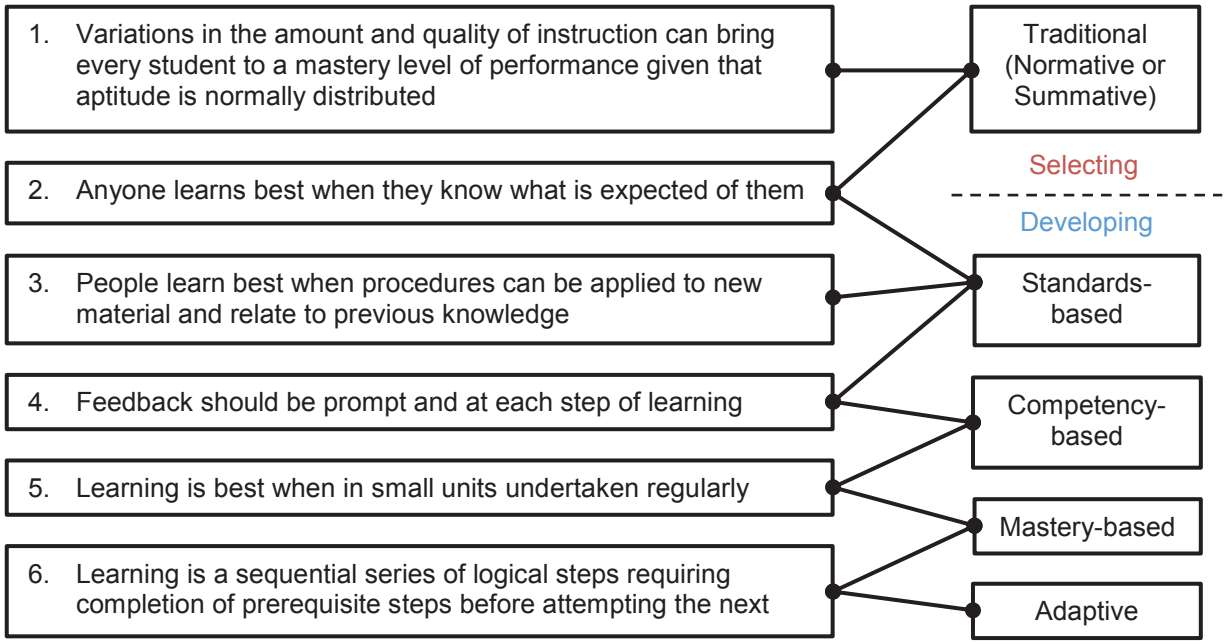

Fig. 1. Mapping of criterion-referenced principles (from Heywood, 2016) and assessment methods found in the literature.

5 A GRADING SYSTEM CONTINUUM

The defining principles of grading systems can be further expounded upon to better understand the degree in which each system attends to these principles and differentiates itself from other grading systems. Fig. 2 places each grading system on a continuum from developing to selecting talent by examining differences in how the grading system impacts sequencing of content, course pace, number of attempts to demonstrate achievement, scale and weight of performance, types of feedback provided, and basis for a final grade.

There is little need to further elaborate on traditional, norm-based grading approaches. The continuum captures the previous discussion, including the primary action of adding up a number of discrete and disparate tasks throughout the semester to determine a student's performance using a 100-point scale or letter grade. LO-based grading approaches have clearly demonstrated variation that has led to subsequent confusion and inconsistent labelling of these systems across instructors and programs. The presented continuum defines each grading system in an effort to alleviate confusions that have emerged due to a lack of a tool to help guide the use and description of these systems.

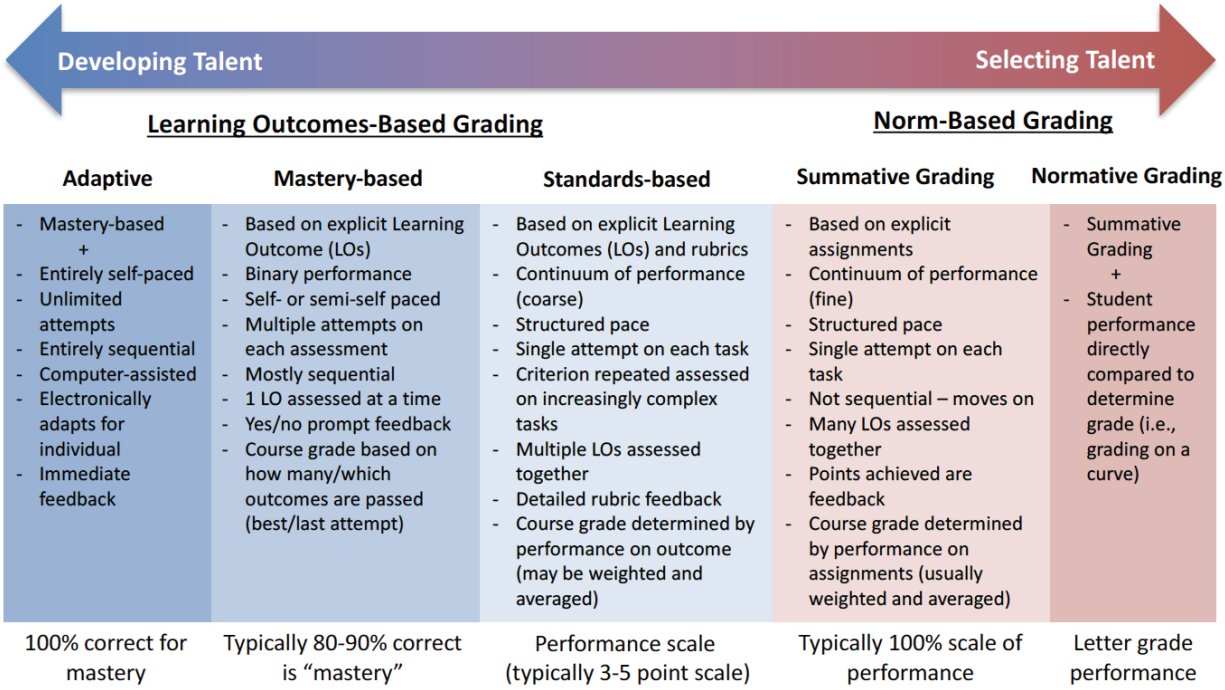

Fig. 2. Continuum of various assessment methods used in engineering higher education.

6 GRADING SYSTEM REFORM

The primary intent of the grading system continuum is to clarify what has become a confusing landscape of grading systems. A secondary goal is to drive grading system reform as a necessity for the future (R.A. Voorhees, 2001). It was estimated in 2015 that over 600 colleges have or are attempting to implement some form of

LO-based grading (Fain, 2015), but it's unclear what has been the basis for such change and how efforts are being benchmarked. An example of a large systematic attempt at grading system reform that can be referenced occurred in primary and secondary education within the United states (US) back in the 1990s. The effort was in response to increasing legislative involvement to set content and performance standards in public education to improve and increase accountability toward monitoring student learning (Marzano, 2011; Reeves, 2002). The primary challenge associated with such a movement is the difficulty overcoming the pervasive mindset that the primary function of grading is differentiating between students rather than assessing a particular student's achievement or competency (Kendall & Marzano, 1996; Guskey, 2011); yet calls for alternative grading approaches in K-12 have risen again with recent concern for accountability in the public education system and online learning (Spencer, 2017).

Adopting a similar approach in higher education should consider the challenges faced in US K-12 education; however, Marzano (2011) and Kendall & Marzano (1996) note that one size does not fit all and that tailoring approaches is difficult to balance against the broadly sweeping standardized criteria set by a state or national panel. These considerations are noted within the context that K-12 and higher education have inherently different structures. Much of the pushback at the K-12 level has arisen from standards that are too exhaustive and do not allow for individual instructors to determine how best to teach the material. Such angst may be inherently ameliorated in higher education because equivalent course-level standards across institutions do not exist. Instructors maintain much more autonomy over their classroom structure and assessment even when designing a course to fit within accreditation standards or institutional goals (Eaton, 2016). This presents an interesting opportunity for courses designed to be "tailor-made to individual schools" (p. 13), which is described as a must for successful implementation (Kendall & Marzano, 1996).

The challenge associated with changing a grading system in such a fashion is that course changes are rarely accompanied by institutional changes, i.e., grades still need to be converted. Additional challenges include a lack of consensus on how to structure courses, what assessment techniques are best, and what actually constitutes a full conversion to an LO-based grading system (Witt et al, 2006). Establishing a consensus and differentiating the variations used by engineering educators will help provide guidelines for those wishing to convert and a standard upon which to compare different systems (JEE).

7 SUMMARY

Higher education engineering instructors have a variety of grading systems to choose from when assessing their students. Traditional, norm-based approaches are still widely used, but alternative LO-based approaches are starting to see greater use in engineering classrooms. The initial overhead to change current practices can be seen as quite large, but the benefits in the long run are worthwhile for students, instructors, and programs (Lee et al., 2018). The presented continuum is a tool that instructors and programs can use to understand the options and help make an informed decision for an individual course or curriculum. It is our role as researchers to help educate our colleagues about grading options using tools like the continuum presented in this paper in order to address greater institutional accountability (A.B. Voorhees, 2001).

8 ACKNOWLEDGMENTS

This work was made possible by a grant from the National Science Foundation (NSF DUE-1503794). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Postman, N. (1992), Technopoly: The Surrender of Culture to Technology, Alfred A. Knopf, New York, NY.
[2] Guskey, T. R. and Pollio, H. R. (2012), Grading systems – school, higher education, Sage Publications, Thousand Oakes, CA.
[3] Muñoz, B. M. A. and Guskey, T. R. (2015), Standards-based grading and reporting will improve education, *Phi Delta Kappan*, Vol. 96, No. 7, pp. 64-68.
[4] Freeman, R. and Lewis, R. (1998), Planning and implementing assessment, Kogan Page, London, U.K.
[5] Huba, M. E. and Freed, J.E. (2000), Learner-centered assessment on college campuses: Shifting the focus from teaching to learning, Allyn & Bacon, Needham Heights, MA.
[6] Morgan, C., Dunn, L., Parry, S. and O'Reilly, M. (2004), The student assessment handbook: New directions in traditional and online assessment. Routledge Falmer, London, U.K.
[7] Stevens, D. D. and Levi, A. J. (2004), Introduction to rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning. Stylus, Sterling, VA.
[8] Suskie, L. (2004), Assessing student learning: A common sense approach. Anker, Boston, MA.
[9] Walvoord, B. E. and Anderson, V. J. (1998), Assessing student learning: A common sense approach, Anker, Boston, MA.
[10] Cross, L. H. and Frary, R. B. (1999), Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Educati*on, Vol. 12, No. 1, pp. 53-72.
[11] Angelo, T.A. and Cross, K.P. (1993), Classroom Assessment Techniques: A Handbook for College Teachers, Jossey-Bass, San Francisco, CA.
[12] Broad, B. (2000), Pulling you hair out: Crises of standardization in communal writing assessment, *Research in the Teaching of English*, Vol. 35, pp. 213-260.
[13] Sadler, D. R. (1987), Specifying and promulgating achievement standards, *Oxford Review of Education*, Vol. 13, No. 2, pp. 191-209.
[14] Sadler, D. R. (2005), Interpretations of criteria-based assessment and grading in higher education, *Assessment & Evaluation in Higher Education*, Vol. 30, No. 2, pp. 175-194.
[15] Sadler, D. R. (2009), Indeterminacy in the use of present criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, Vol. 34, No. 2, pp. 159-179.
[16] Shay, S. B. (2005), The assessment of complex tasks: A double reading. *Studies in Higher Education*, Vol. 30, No. 6, pp. 663-679.
[17] Diamond, R. M. (2008), Designing and assessing courses and curricula, 3rd ed, Jossey-Bass, San Francisco, CA.

[18] Felder, R. and Brent, R. (2003), Designing and teaching courses to satisfy the ABET engineering criteria, *Journal of Engineering Education*, Vol. 92, No. 1, pp. 7-25.
[19] Krumboltz, J. D. and Yeh, C. J. (1996), Competitive grading sabotages good teaching. *Phi Delta Kappan*, Vol. 78, No. 4, pp. 324-326.
[20] Heywood, J. (2014), The evolution of a criterion referenced system of grading for engineering science coursework, Proceedings of the ASEE/IEEE Frontiers in Education Conference, Madrid, Spain.
[21] McIntyre-Hite, L. (2016). A Delphi study of effective practices for developing competency-based learning models in higher education, The Journal of *Competency-Based Education*, Vol. 1, No. 4, pp. 157-166. DOI: 10.1002/cbe2.1029
[22] Burke, J. W. (1989). Competency based education and training. Bristol, PA: Psychology Press, pp. 109-131.
[23] Spady, W. G. (1977). Competency based education: A bandwagon in search of a definition. *Educational Researcher*, Vol. 6, No. 1, pp. 9-14.
[24] Heywood, J. (2016), The assessment of learning in engineering education: Practice and policy, John Wiley & Sons, Inc., Hoboken, NJ.
[25] Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance, *British Journal of Educational Psychology*, Vol. 58, No. 1, pp. 1-14.
[26] Guskey, T. R. (1997), Research needs to link professional development and student learning, *Journal of Staff Development*, Vol.18, pp. 36-41.
[27] Post, S. L. (2014), Standards-based grading in a fluid mechanics course, Proceedings of the American Society for Engineering Education Annual Conference & Exposition, Indianapolis, IN.
[28] Wiggins, G. and McTighe, J. (1998), Understanding by design. ASCD, Alexandria, VA.
[29] Henri, M., Johnson, M. D., and Nepal, B. (2017), A review of competency-based learning: Tools, assessments, and recommendations. *Journal of Engineering Education*, Vol. 106, No. 4, pp. 607-638.
[30] Atwood, S., Siniawski, M., & Carberry, A. (2014). Using standards-based grading in engineering project courses. Proceedings of the American Society for Engineering Education Annual Conference & Exposition, Indianapolis, IN.
[31] Carberry, A.R., Siniawski, M., and Dionisio, J. (2012), Standards-based grading: Preliminary studies to quantify changes in student affective and cognitive behaviors, Proceedings of the ASEE/IEEE Frontiers in Education Conference, Seattle, WA.
[32] Siniawski, M., Carberry, A., and Dionisio, J. (2012), Standards-based grading: An alternative to score-based assessment, Proceedings of the American Society for Engineering Education Pacific Southwest Section Conference, San Luis Obispo, CA.
[33] Hylton, J. B. and Diefes-Dux, H. A. (2016), A standards-based assessment strategy for written exams, Proceedings of the American Society for Engineering Education Annual Conference & Exposition, New Orleans, LA.
[34] Lee, E., Carberry, A. R., Atwood, S. A., Diefes-Dux, H. A., and Siniawski, M. T. (2018), Faculty perceptions before and after implementation of standards-based grading, *Australasian Journal of Engineering Education*, Vol. 23, No. 2, pp. 53-61.

[35] Diefes-Dux, H. A. (2018), Student self-reported use of standards-based grading resources and feedback, *European Journal of Engineering Education*, DOI: 10.1080/03043797.2018.1483896
[36] Post, S. (2017), Standards-based grading in a thermodynamics course, *International Journal of Engineering Pedagogy*, Vol. 7, No. 1, pp. 173-181.
[37] Bekki, J. M., Dalrymple, O. and Butler, C. S. (2012), A mastery-based learning approach for undergraduate engineering programs. Proceedings of the ASEE/IEEE Frontiers in Education Conference, Seattle, WA.
[38] DeGoede, K. M. (2018), Competency based assessment in dynamics. Proceedings of the American Society for Engineering Education Annual Conference & Exposition, Salta Lake City, UT.
[39] Padhye, L. P. and Blumenstein, M. (2017), In-class and asynchronous student response systems: A comparison of student participation and perceived effectiveness. Proceedings of the Australasia Engineering Education Conference, Sydney, Australia.
[40] Voorhees, R. A. (2001). Competency-Based learning models: A necessary future. *New Directions for Institutional Research*, Vol. 110, pp. 5-13.
[41] Fain, P. (2015) Amid competency-based education boom, a meeting to help colleges do it right. Inside Higher Education Available from: https://www.insidehighered.com/news/2015/09/10/amid-competency-based-education-boom-meeting-help-colleges-do-it-right [last accessed 1 August 2019].
[42] Marzano, R. J. (2011), Formative assessment and standards-based grading. Marzano Research Laboratory, Bloomington, IN.
[43] Reeves, D. B. (2003), Making standards work: how to implement standards-based assessments in the classroom, school, and district, Advanced Learning Press, Englewood, CO.
[44] Kendall, J. S. and Marzano, R. J. (1996), Content knowledge: A compendium of standards and benchmarks. for K-12 education. Mid-Continent Regional Educational Laboratory, Aurora, CO.
[45] Guskey, T. R. (2011), Five obstacles to grading reform, *Educational Leadership*, Vol. 69, No. 3, pp. 17–21.
[46] Spencer, K. (2017), A new kind of classroom: No grades, no failing, no hurry. New York Times, MB1.
[47] Witt, H.-J., Alabart, J., Giralt, F., Herrero, J., Vernis, L., and Medir, M. (2006), A competency-based educational model in a chemical engineering school, *International Journal of Engineering Education*, Vol. 22, No. 2, pp. 218–235.
[48] Eaton, J.S. (2016). Accreditation and competency-based education. The *Journal of Competency-Based Education*, Vol. 1, No. 1, pp. 12-16.
[49] Voorhees, A. B. (2001). Creating and implementing competency-based learning models. *New Directions for Institutional Research*, Vol. 110, pp. 83-95.