A novel taxonomy to assess engineering students: the Face-it project

Marica Liotino[1], Monica Fedeli[2], Anja Garone[3], Steffi Knorn[4], Damiano Varagnolo[5], Emanuele Garone[6]

ABSTRACT:

Formally describing and assessing the difficulty of learning and teaching material is important for quality assurance in university teaching, for aligning teaching and learning activities, and for easing communications among stakeholders such as teachers and students. This paper proposes a novel taxonomy to describe and quantify the difficulty levels of exam questions and exercises encountered in engineering-related contexts. This paper also describes the development and piloting processes of the new taxonomy. The proposed taxonomy consists of two dimensions which describe the difficulty in understanding/explaining and using/applying a content unit. The piloting phase included ten purposefully selected experts in the field of control engineering, external to the project, who tested the performance, utility, ease of use, and clarity of the new taxonomy. The results indicate that the users were able to provide consistent and coherent assessments of the difficulty levels of 15 selected exam questions. The paper further discusses suggestions for improvement voiced by the participants to promote an even more consistent and coherent assessment of engineering students' mastery of the subject.

Keywords: assessment, taxonomy, higher education, engineering

Introduction

This study is part of an Erasmus+ project titled "Face It: Fostering Awareness on Program Contents in Higher Education using IT tools", realized by Uppsala University, Université Libre de Bruxelles, Otto-von-Guericke University Magdeburg, Norwegian University of Science and Technology, and the University of Padua. It brings together Engineering and Pedagogy to develop new shared methods for defining, collecting, managing, processing and visualizing program content in association with program learning objectives, teaching-learning activities (TLAs), and intended learning outcomes (ILOs). The focus is on improving the common understanding of what is being taught in courses or programs, what is expected from students, and how the courses in a program are connected. Thus, it includes improving the clarity, efficiency, and effectiveness of the forms of information exchange among teachers, students, and administrative staff. Indeed, the educational community involved reported that in the engineering educational field frustration often arises when discovering that different courses teach the same content, assume prior knowledge that has not been provided yet, or the outcomes as intended by the teachers do not correspond with the learning outcomes perceived by the students. Parts of these problems can be tackled using constructive alignment, a

[1] University of Padua, Italy, e-mail: marica.liotino@phd.unipd.it
[2] University of Padua, Italy, e-mail: monica.fedeli@unipd.it
[3] Vrije Universiteit Brussel, Belgium, e-mail: Anja.Garone@vub.be
[4] Technical University of Berlin, Germany e-mail: steffi.knorn@ovgu.de
[5] Norwegian University of Science and Technology in Trondheim, Norway, e-mail: damiano.varagnolo@ntnu.no
[6] Free University of Brussels, Belgium, e-mail: egarone@ulb.ac.be

curriculum design approach that seeks to optimize the conditions for quality learning, as well as building a coherent learning environment where teaching methods and assessment practices are aligned with teaching objectives (McMahon, & Thakore, 2006). It leverages the following: the definition of the ILOs to describe what students should know and be able to do at the end of the course; the definition of TLAs designed to guide the students towards achieving the ILOs; and the identification of criteria and methods of assessment (Biggs, 2003). In this direction, a curriculum development process started within the Face It project. Firstly, a new taxonomy has been developed, having as primary purpose the definition of a shared and clear lexicon to describe the content of engineering curricula and the related difficulty. To introduce the taxonomy and its development and assessment process are the purpose of this paper.

From a pedagogical perspective, taxonomies have their roots in the curriculum design and development movements and their focus on sequential structuring and objective assessment of learning. Typically, taxonomies stem from behaviorist models of task analysis (analysis of the basic requirements for performing a task) and the construction of learning process feedback systems -originating from the cognitivist framework (Bonaiuti et al., 2017). Indeed, Bloom et al. (1956) first introduced the concept of taxonomy of educational objectives with the aim of reducing the ambiguity of teaching activities, and of organizing in a sequential way the process of assessment. The goal was to identify expected behaviors and the required skills for their achievement. In the last few years, curriculum development processes increasingly considered a student-centered approach (Guerrero-Roldán & Noguera, 2018; Vonderwell & Boboc, 2013), which promotes an ongoing assessment process with different goals. In these processes, assessments are seen as events that are useful for learning, as learning, and of learning; thus, they are a combination of different types of assessment, both formative and summative, which aims to support students' learning, promote their self-regulation, and assess competences (Guerrero-Roldán & Noguera, 2018; Hume & Coll, 2009; Masuku et al., 2021; Vonderwell & Boboc, 2013). Furthermore, growing attention is given to active learning (Cooperstein & Kocevar-Weidinger, 2004; Guerrero-Roldán & Noguera, 2018; Khan et al., 2017; Tabrizi & Rideout, 2017), which implicates "instructional activities involving students in doing things and thinking about what they are doing" (Bonwell & Eison, 1991, p. 5). Active learning "is the process of engaging learners with the topic and each other where they are talking, doing, and creating, together" (Fedeli & Bierema, 2019, p. 30); in this way students construct their learning interacting with the context, involving higher order thinking, and delving into their attitudes and values (Fedeli & Bierema, 2019; Matsushita, 2018).

According to the studied literature, the "Taxonomy of educational objectives" (Bloom et al., 1956) is the most well-known and used taxonomy in Engineering higher education. It consists of three domains (cognitive, affective, and psychomotor) each one divided into categories (Bloom et al., 1956; Krathwohl et al., 1964; Simpson, 1971). The cognitive domain is widespread amongst engineering educators as a framework to describe complexity and higher order thinking (Mead & Bennett, 2009; Stotsky, 2017). Some scholars found the taxonomy useful to design and assess software engineering courses (Britto & Usman, 2015), and to improve the alignment of assessment and learning outcomes in software engineering teaching (Khairuddin & Hashim, 2008). Among the strengths, the extensive analysis of test items, its simplicity, and the distinctness of factors of the cognitive domain have been identified (Fuller et al., 2007). Nevertheless, some critiques have been addressed: It is not suitable for the computing context (Azuma et al., 2004; Masapanta-Carrión & Velázquez-Iturbide, 2018) and does not adequately address the skills and competences needed in engineering (Heywood, 2005, p.28). Other highlighted difficulties concern the differentiation of the cognitive activity involved in each category (Fuller et al, 2007; Masapanta-Carrión, & Velázquez-Iturbide, 2018; Staffas et al., 2020). As a consequence of the above, considering students' cognitive processes can become challenging (Kallia, 2017; Masapanta-Carrión & Velázquez-Iturbide, 2018). Additionally, this taxonomy offers different interpretations (Heywood, 2005; Johnson & Fuller, 2006; Staffas et al., 2020) and overlaps among categories (Fuller et al., 2007) that make some learning goals fit into more than one category (Masapanta-Carrión & Velázquez-Iturbide, 2018; Staffas et al., 2020). There are also disagreements in categorizing knowledge related to higher levels (Azuma et al., 2004; Fuller et al, 2007). In fact, the applicability of these categories to every module has been problematic (Johnson & Fuller, 2006) and therefore not suitable for undergraduate courses (Ardis et al., 2015 p.17). Furthermore, some authors consider the taxonomy not exhaustive, as it neglects operational knowledge (Azuma et al., 2004) and accordingly have proposed the addition of other categories to the taxonomy (Heywood, 2005). Many attempts have been made to standardize Bloom's taxonomy use (Britto & Usman, 2015; Masapanta-Carrión & Velázquez-Iturbide, 2018). Some scholars proposed modifications of the original Bloom's Taxonomy (Azuma et al., 2004; Fuller et al., 2007; Johnson & Fuller, 2006), some others used the revised version proposed by Anderson et al. in 2001 (Amorim et al., 2014; DeMara et al., 2019; Froyd et al., 2012), or a further modification created in 2007 by Marzano and Kendall (Vargas-Mendoza et al., 2018).

Also known in the field is the SOLO (Structure of the Observed Learning Outcome) taxonomy (Biggs & Collis, 1982). According to Biggs and Collis (1982) this taxonomy aims to pay more attention to the authenticity of the evaluation in terms of using "levels that arise naturally in the understanding of the material" instead of a priori ones (Biggs & Collis, 1982, p. 13). The SOLO taxonomy, consisting of five levels of knowledge based on Piaget's stages of cognitive development, aims at capturing adult conceptual development. Intuitiveness and reliability (Stotsky, 2017; Watson et al., 2014), usefulness to analyze student's knowledge (Kallia, 2017; Watson et al., 2014), and the holistic nature (Fuller et al., 2007) are its strengths. On the other hand, imprecision on specific concepts learned (Staffas et al., 2020; Watson et al., 2014) and lack of usage experiences in the field (Fuller et al., 2007) have also been reported. Modifications have been proposed to the original version (Stotsky, 2017) and a priori coding scheme has been adopted to improve its usage (Watson et al., 2014). Other lesser-known taxonomies have also been proposed such us the Taxonomy of Significant learning proposed by Fink in 2003(Man Choi, 2019), or inedited taxonomies (Crawley et al.,2011; Sedelmaier & Landes, 2012).

Based on these considerations, the already existing taxonomies do not meet the educational and assessment need of the Face-It educational community. Consequently, the research group has drafted a new taxonomy that should be intuitive, valid, reliable, and suitable for defining the knowledge and skills required to successfully answer typical questions and exercises in engineering. Therefore, the study is guided by a multiple theoretical framework rooted in curriculum development, which pays attention to the creation of a coherent learning environment in the perspective of constructive alignment and sees in taxonomy a tool that promotes clear information exchange among stakeholders and supports the two just mentioned.

The Proposed taxonomy

Considering existing taxonomies and inspired by the division of engineering knowledge into procedural and conceptual knowledge, the proposed taxonomy assumes the difficulty of a question as measurable along two dimensions: using and explaining. The taxonomy leverages the concept of content unit (CU), with which we mean an atomic unit of knowledge, e.g., electric potential, Rouché-Capelli theorem. In other words, we assume that each question corresponds to an opportune set of CUs that indicate which content the question covers. The taxonomy level of a question should then indicate how difficult the question is (e.g. using 2; explaining 1). In principle, all combinations are possible apart from the level u0, e0.

The Using Dimension

This dimension is dedicated to measuring the increase in difficulty of the skills needed to compute a correct answer, solve a problem, or derive a quantitative result. The levels are:

- Level u0 (short for “using level 0”) are questions that do not require computing a specific output.
- Level u1 are questions that ask explicitly to obtain some quantitative outputs, tell explicitly which CUs should be used to compute the outputs, and tell explicitly how to use these CUs if these can be used in more than one way.
- Level u2 are questions that ask explicitly to obtain some quantitative outputs, only hint at which CUs should be used to compute the outputs, and only hint at how to use these CUs.
- Level u3 are questions that ask explicitly to obtain some quantitative outputs, neither tell nor hint at which CUs should be used to compute the outputs, and neither tell nor hint at how to use the CUs.

The “Explaining” Dimension

This dimension is dedicated to measuring the increasing difficulty of conceptual knowledge and understanding needed to arrive at a correct answer, explain or predict a phenomenon or behavior, or derive a qualitative result. The levels are:

- Level e0, questions that require pure computations without any explanation or reasoning.
- Level e1 are questions that can be answered just through memory recalling operations, such as questions asking the student to define or recall the explicitly mentioned CUs, to recognize the correct keywords or a phrase that defines the mentioned CUs.
- Level e2 are questions that simultaneously clearly mention or hint at both the CUs involved, the technical context, and at least hint at a pre-described or obvious path

to reach the solution; they cannot be answered through only memory recalling operations, because they require also performing cognitive/logical connections among the ingredients above to reach an outcome that is explicitly specified in the question and require logical connections among the ingredients above. Such questions may ask the student to do the following: describe the CUs in their own words, add information to the main points characterizing the CUs, interpret and summarize the CUs, construct a symbolic representation of the CUs, and/or translate the CUs from one form to another, for example through figures or diagrams.

- Level e3 are questions that present at least one of the following features: they do not mention explicitly or hint clearly at all the ingredients needed to answer the question, nor do they hint at a pre-described or obvious path to get the solution, or require the student to choose from multiple possible nontrivial paths to reach a correct solution; they require constructing upon previous knowledge, i.e., performing logical connections beyond what is explicitly mentioned in the exercise, and thus require extrapolating information to correctly predict and/or generalize concepts, consequences and/or phenomena in other contexts and/or outside the subject area. Hence, such questions may ask the student to do the following: recognize some relationship like similarities, differences and cause-effect relationships between the ingredients in the question and some nonexplicitly mentioned CUs; identify errors in the presentation or use of some explicitly mentioned CUs; solve questions that require the application of CUs in some specified situations/contexts but at the same time provide only incomplete information, and thus require the student do logical connections beyond what is explicitly mentioned; recognize some organizational principles involving the mentioned CUs; consider some trans-disciplinary aspects of the mentioned CUs; require the student to perform analyses, or form opinions, estimates or predictions that necessarily involve what is beyond what is explicitly mentioned in the exercise.

Methodology

Not having found a satisfactory and validated process of taxonomy development and validation in the literature, we studied the processes previously followed by scholars for the validation of their taxonomies and we took as reference the process for the scale validation of Boateng et al. (2018). Firstly, we identified the domain and generated the items. A content validity assessment followed. According to Boateng et al. (2018) this is best done through the combination of external expert judges and target-population judges; therefore, ten participants with these features were recruited via email. They belonged to the professional network of the authors and were not involved in the crafting of the taxonomy. Ten meetings (one for each participant) were held using a video-call platform (Zoom) between October 27, 2020, and November 20, 2020, and recorded to facilitate later analysis. The participants (one female, nine male) were academic members (two full professors, two associate professors, two assistant professors, one senior researcher and one postdoc) and other professionals from industry, working in Europe (n=7) and Northern America (n=3) in the same scientific area (Systems and Control) and with teaching experience ranging from 0 (meaning at most limited to assisting with preparing exams) to 30 years. During the meetings, the participants were asked to read the manual created to explain how to use the taxonomy. Secondly, they assessed the taxonomical level of a set of 15 questions according to the directions written in the manual. The aim of this exercise was to measure whether participants were using the taxonomy consistently. Each meeting ended with a semi structured interview (Trinchero, 2002) to explore the following:

- Clarity, especially regarding the lexicon, structure and purpose, to understand if the taxonomy is described well and unambiguous (Boateng et al., 2018; Mountrouidou et al., 2019; Wolever et al., 2020).
- Exhaustiveness, in terms of completeness (Huff et al., 1984; Mountrouidou et al., 2019; Tett et al., 2000), i.e., being composed by all the dimensions and categories needed to categorize the difficulty of exercises.
- Effectiveness, i.e., if it achieves the established objectives (Alvino et al., 2006; Bezzi, 2007; Pozzoli & Manetti, 2011), that in this case is the classification of the difficulty of exercises and the relative labelling.
- Relevance, in terms of usefulness for the purpose of assessing the difficulty of exercises, and usefulness in the assessment process of teaching (Boateng et al., 2018; Devon, 2007; Huff et al., 1984; Valentijn et al., 2015; Wolever et al., 2020).
- Distinctness between levels (Spangler & Kreulen, 2002), i.e., "whether the categories are mutually exclusive" (Huff et al., 1984, p.31), whether the exercise is "uniquely represent[ed]" in each dimension (Tett et al., 2000, p. 219) and whether each category is decoupled from others (Mountrouidou et al., 2019, p.7).

Critical issues and suggestions were also collected. The digital transcripts of interviews were analyzed with AtlasTi.08 software, a CAQDAS (Computer Assisted Qualitative Data Analysis Software) that supported the text analysis. A statistical analysis was conducted on Excel for the quantitative data gained from the taxonomic assessment of questions.

Results and Discussion

The results are described in two distinct parts, one dedicated to the quantitative analysis of the measured assessment patterns, and one to the qualitative analysis of the user perceptions and recommendations.

Quantitative Results of Taxonomic Assessment of Questions

Out of 15 questions, only in three questions did more than 30% of the participants disagree with the levels assigned on average by peers. This shows a fairly good level of convergence to the same general assessments of the scale. In the average there is a convergence around 80% to the mode level and the adjacent ones (more precisely 80.5% for the u dimension and 81.0% for e). The convergence to a single class is also fairly good (on average 63.1% converged to the same u level and 51.6% to the same e level), especially if one considers that 10 items have an even distribution between two classes. In several questions there exist some outliers (e.g., Q1, Q2, and Q12). By triangulating the data with the interviews, we realized that most of the votes marked as outliers tend to come from participants whose choice was often due to misinterpretations of the taxonomy manual. Thus, further work must be done to enhance the manual clarity. Overall, the quantitative data show encouraging results, which could be improved with some fine tuning on the manual and wording of the taxonomy levels. We also learned the importance of selecting questions/exercises that are clearly understandable by people coming from different schools and institutional cultures.

Qualitative Analysis of Interviews

The first considered factor is the perceived clarity of the taxonomy. All the 10 participants declared that the purpose of the taxonomy is clear. However, eight of them perceived its structure as clear, and seven perceived its lexicon to be clear. Relative to this, three participants identified some words used in the manual as critical issues. Furthermore, two participants explicitly mentioned that the explaining dimension is noticeably less intuitive than the using one. We believe that this is likely connected to some issues on the distinctness between levels.

The perceived distinctness between the levels is not as accurate as hoped, in fact the participants reported to be unsure in labelling choice while compiling the assessment. The using dimension is associated with noticeably less doubts than the explaining one, and e2 vs. e3 seems to offer the most fleeting discerning boundary. A shared feeling is that the differentiation between levels seems clear on paper (i.e., when reading the manual), but then this clarity diminishes when trying to apply the taxonomy. This calls for adding more examples in the next rewriting of the manual and explaining the levels better.

The next considered factor is the perceived efficacy of the taxonomy. Every participant agreed on the usefulness in labelling exercises; two of them expressed doubts about whether classifying difficulties can be performed in a purely objective way at all. In our stance, it is unlikely that a taxonomy removes all subjectivity effects on indexing difficulty levels, however, having some explicit guide to follow can contribute to reducing that. As for the perceived exhaustiveness of the taxonomy seven participants found that there were dimensions missing (e.g., the time dimension, in the sense of indicating how much time will be required for an average student to solve it, and the complexity dimension, in terms of measuring how tedious and challenging the exercise is). Understanding which and if adding dimensions is needed is still an open question and our current research focus. Another considered factor is the perceived relevance of the taxonomy: seven participants said that already in this form, the proposed taxonomy seems useful for their teaching, especially as a tool for aligning the expectations with the students and colleagues, on top of sharing material within the community.

Finally, perceived strengths and weaknesses were analyzed. Among reported weaknesses, the current taxonomy does not promote enough distinctness between the various levels, and it is insufficiently exhaustive, i.e., it lacks dimensions to capture the various shades of difficulty of various exercises. Two persons mentioned that already the existence of this taxonomy is a strength, i.e., already having something that covers a perceived gap is a strength. Moreover, the participants mentioned expected benefits of taxonomy employment as strengths e.g., the possibility of aligning expectations with the various stakeholders on top of exchanging teaching material, the possibility of checking the consistency of the exams' difficulty levels across the years and promoting teachers' reflections on the exercises.

Conclusions

The focus of the paper is on introducing a new taxonomy the purpose of which is to enable an objective indexing of automatic-control-related assessment material. Item generation and content validity assessment has been presented. In the latter, ten people were recruited in an indexing exercise accompanied by a semi-structured interview. Both quantitative and qualitative analyses were carried out, highlighting that ill-posed questions are evidently associated with higher-than-normal spreads of the indexing, the taxonomy is still incomplete from a dimensional point of view and has issues on differentiating levels of difficulties. Currently our efforts are in reformulating taxonomy and exercises to address the issues encountered, as well as planning the content validity assessment of the second version. Future work may try to subject the taxonomy to the evaluation of a statistically relevant sample and to introduce the taxonomy in other areas and disciplines. Finally, we believe that a pedagogical research study design could give evidence on how the integration of this taxonomy in the teaching-learning process can support students' learning, giving them a more accurate feedback and assessment on their learning processes and outcomes.

References

Alvino, S., Busetti, E., Forcheri, P., & Ierardi, M. G. (2006). "Networked learning and reusable teaching resources". *Italian Journal of Educational Technology* (IJET), 14(3), 30-38. https://ijet.itd.cnr.it/article/view/403

Amorim, G. F., Balestrassi, P. P., Paiva, A. P., & Gottzandt, I. S. R. (2014) A didactic activity for introducing design and optimization of experiments assisted by revised Bloom's taxonomy. *International Journal of Higher Education*, 3(4), 12–23. https://doi.org/10.5430/ijhe.v3n4p12

Ardis, M., Budgen, D., Hislop, G. W., Offutt, J., Sebern, M., & Visser, W. (2015). SE 2014: curriculum guidelines for undergraduate degree programs in software engineering. *Computer*, 48(11), 106–109. https://doi.org/10.1109/MC.2015.345

Azuma, M., Coallier, F., & Garbajosa, J. (2004) How to apply the Bloom taxonomy to software engineering. *Eleventh Annual International Workshop on Software Technology and Engineering Practice*. doi:10.1109/step.2003.13

Bezzi, C. (2007). Cos'è la valutazione: un'introduzione ai concetti, le parole chiave e i problemi metodologici. Franco Angeli. https://books.google.it/books?id=tr4Lo3I3lf4C

Biggs, J. (2003). Aligning teaching for constructing learning. *Education*, 94(11), 112106. https://doi.org/10.1063/1.3100776

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO Taxonomy, Structure of the Observed Learning Outcome*. London, United Kingdom: Academic Press, Inc.

Bloom, B. S., Engelhart, 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

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health*, 6(June), 1–18. https://doi.org/10.3389/fpubh.2018.00149

Bonaiuti, G., Calvani, A., & Ranieri, M. (2017). Fondamenti di didattica: Teoria e prassi dei dispositivi formativi. http://www.carocci.it/index.php?option=com_carocci&task=schedalibro&Itemid=72&isbn=9788843042487

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom. 1991 ASHE-ERIC Higher Education Reports*. ERIC Clearinghouse on Higher Education, The George Washington University.

Britto, R., & Usman, M. (2015) Bloom's taxonomy in software engineering education: A systematic mapping study. *Proceedings - Frontiers in Education Conference, FIE, 2015(October)*. https://doi.org/10.1109/FIE.2015.7344084

Cooperstein, S. E., & Kocevar-Weidinger, E. (2004). Beyond active learning: A constructivist approach to learning. *Reference services review*.

Crawley, E. F., & Lucas, W. a. (2011). The CDIO Syllabus v2.0 An updated statement of goals for engineering education. *Engineering Education*, 24, 1–4. http://files.conferencemanager.dk/medialibrary/59856d54-6d1c-4deb-ac21-0dd98ccd0470/images/CrawleyEtAlCDIOSyllabus2.0Paper_17June2011.pdf

DeMara, R. F., Tian, T., & Howard, W. (2019) Engineering assessment strata: A layered approach to evaluation spanning Bloom's taxonomy of learning. *Education and Information Technologies*, 24(2), 1147–1171. https://doi.org/10.1007/s10639-018-9812-5

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., ... & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing scholarship*, *39*(2), 155-164.

Fedeli, M., & Bierema, L. L. (2019). *Connecting Adult Learning and Knowledge Management*. Springer.

Fink, L. D. (2003). *Creating significant learning experiences. An Integrated Approach to Designing College Courses*. Jossey-Bass.

Froyd, J. E., Wankat, P. C., & Smith, K. A. (2012) Five Major Shifts in 100 Years of Engineering Education. *Proceedings of the IEEE, 100(Special Centennial Issue),* 1344-1360. doi:10.1109/jproc.2012.2190167

Fuller, U., Johnson, C. G., Ahoniemi, T., Cukierman, D., Hernán-Losada, I., Jackova, J., Lahtinen, E., Lewis, T. L., Thompson, D. M., Riedesel, C., & Thompson, E. (2007) Developing a computer science-specific learning taxonomy. *ACM SIGCSE Bulletin*, 39(4), 152–170. https://doi.org/10.1145/1345375.1345438

Guerrero-Roldán, A. E., & Noguera, I. (2018). A model for aligning assessment with competences and learning activities in online courses. *The Internet and Higher Education*, *38*, 36-46.

Heywood, J. (2005). *Engineering Education: Research and Development in Curriculum and Instruction* (1st ed.). Wiley-IEEE Press.

Huff, S. L., Rivard, S., Grindlay, A., & Suttie, I. P. (1984). An empirical study of decision support systems. *INFOR: Information Systems and Operational Research*, 22(1), 21–39.

Hume, A., & Coll, R. K. (2009). Assessment of learning, for learning, and as learning: New Zealand case studies. *Assessment in Education: Principles, Policy & Practice*, *16*(3), 269-290.

Johnson, C. G., & Fuller, U. (2006) Is Bloom's taxonomy appropriate for computer science? *ACM International Conference Proceeding Series*, 276, 120–123. https://doi.org/10.1145/1315803.1315825

Kallia, M. (2017) Assessment in Computer Science courses: A literature review. *Royal Society*, 1–60.

Khairuddin, N. N., & Hashim, K. (2008) Application of Bloom's taxonomy in software engineering assessments. *Proceedings of the 8th Conference on Applied Computer Science*, 66–69. http://portal.acm.org/citation.cfm?id=1504034.1504048

Khan, A., Egbue, O., Palkie, B., & Madden, J. (2017). Active learning: Engaging students to maximize learning in an online course. *Electronic Journal of E-Learning*, *15*(2), pp107-115.

Krathwohl, D. R., Bloom, B. S.& Masia, B. B. (1964). *Taxonomy of educational objectives: The classification of educational goals*; Handbook II: The Affective Domain. New York: David McKay.

Man Choi, E. (2019) Software engineering education for significant learning experience. *International Journal of Information and Education Technology*, 9(12), 862–867. https://doi.org/10.18178/ijiet.2019.9.12.1318

Masapanta-Carrión, S., & Velázquez-Iturbide, J. Á. (2018) A systematic review of the use of Bloom's taxonomy in computer science education. *SIGCSE 2018 - Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 2018-January, 441–446. https://doi.org/10.1145/3159450.3159491

Masuku, M. M., Jili, N. N., & Sabela, P. T. (2021). Assessment as a pedagogy and measuring tool in promoting deep learning in institutions of higher learning. *International Journal of Higher Education*, *10*(2), 274-283.

Matsushita K. (2018) an invitation to deep active learning. In: Matsushita K. (Ed.s) Deep Active Learning. Springer, Singapore. https://doi.org/10.1007/978-981-10-5660-4_2

McMahon, T., & Thakore, H. (2006). Achieving constructive alignment: Putting outcomes first. *Quality of Higher Education*, *3*, 10-19.

Mead, P. F., & Bennett, M. M. (2009). Practical framework for Bloom's based teaching and assessment of engineering outcomes. *Optics InfoBase Conference Papers*. https://doi.org/10.1117/12.2208044

Mountrouidou, X., Billings, B., & Mejia-Ricart, L. (2019). Not just another Internet of Things taxonomy: A method for validation of taxonomies. *Internet of Things*, 6(2019), 100049. https://doi.org/10.1016/j.iot.2019.03.003

Pozzoli M., Manetti G. (2011). *Pianificazione e controllo per le aziende non profit. Valutazioni di efficacia, efficienza, economicità ed impatto socio-economico.* FrancoAngeli, Milano.

Sedelmaier, Y., & Landes, D. (2012) A research agenda for identifying and developing required competencies in software engineering. *2012 15th International Conference on Interactive Collaborative Learning, ICL 2012*, 01, 16–20. https://doi.org/10.1109/ICL.2012.6402195

Simpson, E. J. (1971). Educational objectives in the psychomotor domain. *Behavioral Objectives in Curriculum Development: Selected Readings and Bibliography*, 60(2), 1–35. https://files.eric.ed.gov/fulltext/ED010368.pdf

Spangler, S., & Kreulen, J. (2002). Interactive methods for taxonomy editing and validation. *International Conference on Information and Knowledge Management, Proceedings*, 665–668. https://doi.org/10.1145/584792.584913

Staffas, K., Knorn, S., de Carvalho Guerra, A. O. P., Varagnolo, D., & Teixeira, A. (2020). Using different taxonomies to formulate learning outcomes to innovate engineering curriculum towards PBL: perspectives from engineering educators. In *Educate for the future: PBL, Sustainability and Digitalisation 2020* (pp. 310-320). Aalborg Universitetsforlag.

Stotsky, A. (2017). Modified SOLO Taxonomy Model for Constructive Alignment in Automatic Control & Signal Processing Education.

Tabrizi, S., & Rideout, G. (2017). Active learning: Using Bloom's taxonomy to support critical pedagogy. *International Journal for Cross-Disciplinary Subjects in Education (IJCDSE)*, *8*(3), 3202-3209.

Tett, R., Guterman, A., Bleier, A., & Murphy, P. (2000). Development and content validation managerial competence. *Human Performance*, 13(3), 205–251

Trinchero, R. (2002). *Manuale di ricerca educativa*. FrancoAngeli.

Valentijn, P. P., Vrijhoef, H. J. M., Ruwaard, D., Boesveld, I., Arends, R. Y., & Bruijnzeels, M. A. (2015). Towards an international taxonomy of integrated primary care: A Delphi consensus approach. *BMC Family Practice*, 16(1). https://doi.org/10.1186/s12875-015-0278-x

Vargas-Mendoza, L., Gallardo-Córdova, K. E., & Castillo-Díaz, S. (2018) Performance and authentic assessment in a mechanical engineering course. *Global Journal of Engineering Education*, 20(1), 30–38.

Vonderwell, S. K., & Boboc, M. (2013). Promoting formative assessment in online teaching and learning. *TechTrends*, 57(4), 22-27.

Watson, M. K., Pelkey, J., Rodgers, M. O., & Noyes, C. R. (2014) Exploring student sustainability knowledge using the structure of observed learning outcomes (SOLO) taxonomy. *ASEE Annual Conference and Exposition, Conference Proceedings*.

Wolever, R. Q., Kahn, J. A., Davis, J., Shields, D., & Schoenberg, P. L. A. (2020). Introducing the eMCCTM: A validated taxonomy to advance targeted application of mindfulness skills. *Mindfulness*, 11(3), 698–708. https://doi.org/10.1007/s12671-019-01280-x