Using principles of authentic assessment to redesign written examinations and tests.

Verónica Villarroel[a], David Boud[b], Susan Bloxham[c], Daniela Bruna[*a], Carola Bruna[d]

[a]Center for Research and Improvement of Education (CIME), Faculty of Psychology, Universidad del Desarrollo, Chile; [b]Deakin University, University of Technology Sydney, Australia and Middlesex University, UK; [c]University of Cumbria, UK; [d]Universidad de Concepción, Chile

Abstract

Tests and examinations are widely used internationally. Despite their pervasiveness, they tend to measure lower order thinking skills in a decontextualized manner at a time when the literature frequently argues for the benefits of a richer, authentic approach to assessment. The focus of this paper is to improve authenticity in test assessment methods through promoting realism, cognitive challenge and evaluative judgement during the planning, administering and following up of assessment tasks. The article builds on a systematic literature review, in which the main principles of authentic assessment were outlined. In this paper, we posit how these principles can be implemented through the three chronological phases of the assessment process: before, during and after the act of assessment.

Key words: assessment, authentic assessment, authenticity, testing culture.

---

[*] Corresponding author. Email: dbrunaj@udd.cl

Introduction

Tests focused in knowledge reproduction are widely used at universities (Ghosh, Bowles, Ranmuthugala & Brooks, 2017). There is a strong testing culture in South America (Martínez-Rizo & Mercado, 2015), South East Asia (Gitanjali, 2016) and the Middle East (Mahmoud, 2014), as well as in other university systems (Lesage, Valcke, & Sabbe, 2013). Proponents of testing claim that they reduce plagiarism (Richardson, 2015), increase reliability (McColongue, 2012) and are easy to correct (McCabe, & O'Connor, 2014). These are closed-book tests, in which students are not allowed to bring materials or refer to a textbook. They are administered in controlled conditions as timed unseen tests, in which an invigilator is present to ensure students do not cheat (Hinton & Higson, 2017). Through this process, students tend to become passive learners (Altay, 2014) that memorize content rather than comprehending it (Flores, Veiga-Simao, Barros & Pereira, 2015).

Why is the emphasis on memorising information a problem? Remembering is the lowest level of knowledge assessment (Anderson & Krathwohl, 2001) and students quickly forget what they memorize (Rawson, Dunlosky, & Sciartelli, 2013). In addition, students understand that learning is only mechanically repeating data and information (Schell & Porter, 2018). Instead, when students use higher-order cognitive skills to respond to an assessment, such as concluding, designing or evaluating, they gain a deeper understanding (Entwistle, 2009), and show better stability in remembering what was learned (Rawson, et al., 2013). Although there are differences between disciplines, memorization is not the ultimate learning goal in any subject, and memorisation ill-equips

students for the complex demands of life and work they face on graduation. The achievement of deep learning may require progressively advancing towards it, incorporating memory, analysis and transfer , in different weightings, until students become familiar with the cognitive complexity required.

Assessment is critical in the learning process (Kearney, Perkins & Kennedy-Clark, 2015) because it creates a backwash effect on teaching and learning activities (Watkins, Dahlin, & Ekholm, 2005). It prompts opportunities for students to practice higher order thinking skills. Anticipation of assessment has a strong influence on what and how learners study, frames what students do (Boud, 2010), and drives the learning process (Vu & Dall'Alba, 2014). As a result, assessment has been reported as the most effective way to improve of students´ achievement quality (Edström, 2008). When done poorly, it can have the opposite effect.

Why it is necessary to incorporate authenticity in assessment?

To become a good professional, it is not only necessary to master the knowledge and technical skills of the discipline (Guzzomi, Male & Miller, 2015). Other competencies are also required, such as critical thinking and problem solving, decision-making, communication, collaboration and innovation (Partnership for 21st Century Skills, 2010). It is difficult for tests in themselves to promote a deep approach to learning, that requires the construction of knowledge, reflection and collaborative work, which limits the achievement of central objectives of higher education (Endedijk, & Vermunt, 2013).

Improving the assessment process can provide effective support for the development of the skills graduates need today (Medland, 2016). One approach for making this transition is to follow the principles of authentic assessment (Biggs & Tang, 2011). Authentic

assessment is a way to relate learning and work, creating a correspondence between what is assessed in the university and what graduates do in settings in the outside world (Neely & Tucker, 2012). It has an impact on the quality and depth of learning achieved by the student and the development of higher-order cognitive skills (Ashford-Rowe, Herrington, & Brown, 2014). It can support students' growth in personal confidence (Martinez, O´Brien, Roberts, & Whyte, 2018) and autonomous practice (Raymond, Homer, Smith & Gray, 2012). Moreover, it can improve academic engagement (Kearney & Perkins, 2014), motivation (Nicol, Thomson & Breslin, 2014), self-regulation (Ling Lau, 2013), and metacognition (Vanaki & Memarian, 2009).

Method

The purpose of this article is to explore how the advantages of authenticity in assessment can be applied within the 'testing' approach to assessment, as described above. In this way it acknowledges the need to improve rather than reject test methods, given their dominant use in many higher education systems. It explores how the principles of authenticity can be incorporated through the three chronological phases of the assessment process: before, during and after the act of assessment in written tests.

The article focuses on the second part of a two-stage project. In stage 1, a systematic review of the literature was undertaken to identify principles of authentic assessment and in stage 2, reported here, the authors undertook an exploratory application of the principles to a testing environment, identifying illustrative, and exemplar questions. In stage 1, Villarroel, Bloxham, Bruna, Bruna, & Herrera-Seda (2018) carried out a systematic review of 125 articles on authentic assessment published between 1988 and 2017 to identify its main characteristics as encapsulated in the literature. Thirteen central

characteristics were identified, which were grouped into three dimensions that constitute the core of the construct: realism, cognitive challenge and authentic evaluative judgement. Realism is the first principle that distinguishes authentic assessment (Bosco & Ferns, 2014), understood as representing something that might be encountered in the world beyond university. The second principle represents cognitive challenge whereby students use higher-order cognitive skills related to using, modifying, or rebuilding knowledge into something new (Thornburn, 2008). Thirdly, evaluative judgement is a necessary capability of graduates to make decisions about the quality of work of oneself and others. It allows students to anticipate, monitor and improve the quality of their work and that of others (Tai, Ajjawi, Boud, Dawson, P & Panadero, 2017).

Stage 2 adopts a more expository approach. It attempts to posit how the three dimensions of authenticity (realism, cognitive challenge and authentic evaluative judgement) can be applied in a 'testing' assessment environment. To this end, each dimension was mapped against some phases of assessment cycle (and its elements), using a chronological sequence presented in figure 1:
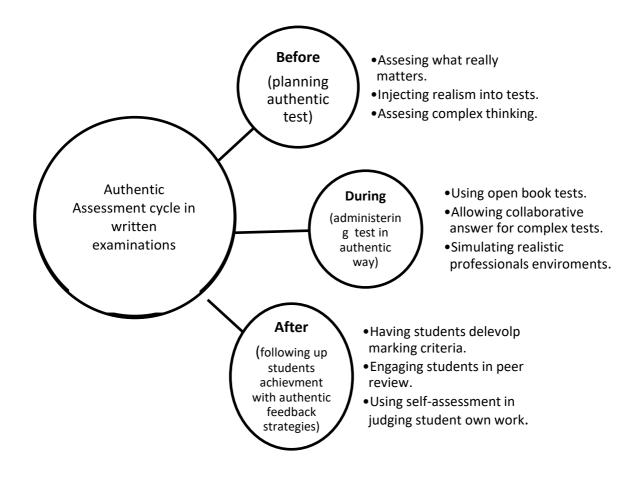
Figure 1. Phases and elements of the implementation of the authentic assessment cycle in written examinations.

In each phase of this chronological framework, we apply the conceptual description of each dimension of authenticity (realism, cognitive challenge and authentic evaluative judgement) to concrete aspects of assessment design, drawing on relevant research literature to support the arguments. Illustrative examples of the principles in practice are provided. Whilst the exploration is not exhaustive in either scope or reference to related studies, it is original in theorising a range of ways in which testing methods can better

reflect the essential features of authentic assessment.

Authenticity in the assessment cycle

I.- Before: planning authentic tests.

1. - Assessing what really matters. In authentic assessment, the validity of what is measured is fundamental. To facilitate the adequate selection of the content, we propose three sources: the graduate profile, course learning outcomes and professional requirements, where they exist. These three elements improve the potential for "realism" in the assessment.

Graduate profile

The graduate profile represents the competences or learning standards that all graduates need to demonstrate once they finish their studies. These are often articulated at an institutional level. This set of general standards (variously called transferable skills, generic attributes) enables course designers to determine how their course will contribute to this profile and ensure that assessment will be orientated to measure pertinent goals (Hart, Hammer, Collins & Chardon, 2011). *How does each subject connect and contribute to achieving the competences of the graduate profile?*

Course learning outcomes

In any educational process, learning outcomes must be assessed, so this is not something exclusive to authentic assessment. However, it is necessary to emphasize the importance that educators ask themselves, what can students do by the end of each course and course

unit. Authentic assessment can be generated by the use of a backward design methodology (Wiggins & McTighe, 2006) which analyses course learning outcomes and identifies which assessments are necessary for it to be claimed that students have met them all.

Professional requirements

Courses which lead to professions also have professional practice requirements. These includes what competences are needed for good professional performance. It is necessary to incorporate ways of assessing competences which will allow students to face typical problems in professional work (Maxwell, 2012). *Do the capabilities acquired in the course allow graduates to respond to the problems or functions needed by the profession?*

An example is shown below of an item from a third year undergraduate course in Speaking and Hearing Therapy, "Disorders and Intervention of the Swallowing"::

Graduation profile:

Solve speaking and hearing problems systematically drawing on evidence and relevant knowledge.

Course learning outcomes:

Identify therapeutic objectives based on the analysis of clinical cases of patients.

Professional requirements:

Plan a functional and neuromuscular evaluation of the phonoarticulatory organs with patients in different stages of the life cycle.

Test item:

In the functional and neuromuscular evaluation of the phonoarticulatory organs of a 6-year-old girl, it was observed that she presents difficulty carrying food from the vestibule to the occlusal face of the molars, therefore she carries the food to the occlusal face with her finger. To determine the therapeutic objectives of your phonoaudiological therapy, you must start, by identifying the muscle that needs rehabilitation to achieve a better functioning in this case. From the following options, select the correct one:

a) Masetero

b) Lateral Pterygoid

c) Buccinator

d) Temporary

2. - *Injecting realism into tests.* Realism can be accomplished by presenting a real context that describes and delivers a frame in which a problem is to be solved. Items can be drafted with rich context simulating real-work situations that function as a proxy for professional performance even when the course does not include assessment in a professional setting. The information presented in the context may show more than one perspective of a phenomenon or create limits or restrictions that must be considered in responding to the problem.

It is not easy to create good contexts. It is a common occurence that questions can be answered without analysing the context. In these, the context constitutes an ornament or a frame which does not have information needed to solve the question. Villarroel et

al. (2018) showed that 47% of 4401 test items in 6 undergraduate programs presented a context. However, in 73% of them, the information within this context was not needed to answer the question. An example of an "ornament" context followed by a well-constructed context in a biology course on the concept of *autophagy,* is presented:

Ornament context

The Nobel Prize in Medicine was assigned to the biologist Oshumi for his discoveries of a process called autophagy. Describe the autophagy process and comment on its implications for health.

Well-constructed context

Andrea and Luis are parents for the first time. Andrea had a complication so she had to be taken to the operating room immediately after delivery. Because of this, she has not been able to breastfeed the baby. The father is very worried, despite the fact that the doctor has told him that this is not a problem for the baby if s/he does not receive nutrients from outside sources during the first hours of life.

- Explain to the father the biological mechanism that allows the baby to support its metabolic requirements.

- After being breastfed, analyse the changes to the metabolism of this newborn.

3. - Assessing complex thinking. It is possible to identify three thinking skill levels (Anderson & Krathwohl, 2001),. The first is related to memory skills (recognition or understanding); the second involves analytical skills for information management (comparing, relating, contrasting, interpreting); and the third compromises transfer skills (judging, deciding, criticizing, suggesting, designing, innovating). Authentic assessment

privileges the measurement of transfer skills, where the emphasis is on "why" students learn that content (Avery, Freeman, & Carmichael, 2012), which corresponds to the "cognitive challenge" principle. This principle seeks that students use knowledge for something, either to "manage" it by performing cognitive activities related to analysis, comparison or solving a problem. Alternatively, they use it to display a professional performance that involves high-order skills, such as evaluating, designing or criticizing. An example of the three levels from neuroscience is described below:

Level 1: Memory Skills

Guillermo had a car accident. His frontal dorsal lateral cortex and ventral hypothalamus were destroyed. Draw and label the sagittal section of the brain, labelling at least 10 damaged structures.

Level 2: Analytical Skills

Guillermo had a car accident. He has damaged structures of the cerebral cortex. The mother listens to the doctor state: "it is necessary to administer, externally, substances such as: insulin, dopamine, leptin, peptides ... to regulate it". Infer the areas of the cerebral cortex that have been damaged, based on the medical indications.

Level 3: Transfer Skills

Guillermo had a car accident. His frontal dorsal lateral cortex and ventral hypothalamus were destroyed. Evaluate severity, explaining three possible consequences according to the damaged structures. Also, suggest one strategy that allows you to improve the quality of his future life, compensating for the effects of the accident.

Multiple-choice questions can be designed in an authentic way (Douglas, Wilson and Ennis, 2012) if they require students to undertake decision-making or problem-solving in a contextualized situation and to justify the option chosen through constructed responses. This new format is more complex and students will take longer. They may score lower because they are not used to these demands. In particular, students who may have learned that success is obtained through memorization (Jensen, McDaniel, Woodard & Kummer, 2014). Students may need to be aided in making such a transition.

II.- During: administering tests.

Sitting for a formal test is a stressful event, uncommon in the world outside educational institutions (Brown, Bull, & Pendlebury, 2013). Tests induce anxiety, affecting self-esteem and self-perceptions as learners, especially if they have previously had bad experiences (Harlen, 2005). In contrast, assessment practices, such as problem-based assignments or project work, are perceived by students to be fairer and more effective (Pereira, Flores & Barros, 2017). How can tests include these benefits of performance-based tasks? Three strategies are proposed that respond to the principle of "realism", because they link the assessment situation with the external world:

1.- Using open-book tests. Students report feeling less anxious and more confident on open-book tests (Betts, Elder, Hartley, & Trueman, 2009). And, in addition, cognitive sciences propose that human cognition is extended beyond the individual mind, encompassing other people, symbolic meanings, environment and artefacts. A mind limited only to what we can remember at a certain time, is not a good preparation for modern life, especially, considering that in workplaces there is access to internet, books

and other people to fulfil tasks.

2.- Allowing collaborative answers for complex tests. Learning is built together with others and in interaction. The concept of the zone of proximal development explains the difference between the individual performance in a given task and the performance achieved when the same task is carried out with someone more capable (Wass, Harland, & Mercer, 2011). Consequently, students with low individual marks obtain higher marks in group tests, also displaying more active learning than in individual tests (Almond, 2009). The level of commitment between members is a factor in achieving a high-performance and learning gains (Johnson & Johnson, 2009). Forming small groups and offering a sufficiently complex task that requires dialogue and discussion can help promote this (Davies, 2009). Students learn more from having to argue for what they thought was the right answer and listening to others' reasons (Zhang, Ding, & Mazur, 2017)

3.- Simulating realistic professional environments. Authentic assessment can have positive outcomes on student engagement and motivation in the learning process (Nicol et al., 2014). It is likely that one reason students perform better is that such tasks help develop their professional identity (Huxham, Campbell, & Westwood, 2012). Therefore, it is important that the conduct of tests emulates workplace´s conditions, for example: sending the test via e-mail and requesting students send their answers in the same way at a stipulated time (O´Moore, & Baldock, 2007) and responding to the test on their laptop in the classroom (not only paper and pencil tests). The following is an example of authentic test administration:

The written test is performed in pairs with open books. The case is sent the previous

day by e-mail, without the associated questions. Then, the questions are delivered in the classroom and students can also work outside of it with their materials.

III.- After: following up.

Feedback is important in any assessment, being one of the most powerful influences on students´ learning (Hattie & Timperley, 2007). To make a feedback process authentic it is necessary to include evaluative judgement activities that prepare students for what they will have to do in the world beyond higher education (Tai et al, 2017), that is identify how to judge good work and apply this to their own work and that of others. It helps them to achieve knowledge, skills and predispositions that underpin lifelong learning activities, promoting the development of autonomy (Carter, Sidebotham, Creedy, Fenwick, & Gamble, 2015) and reflective practice (Tait-McCutcheon, Drake, & Sherley, 2011). Three strategies are proposed:

1.- Having students develop marking criteria. Students can jointly construct criteria for marking using their own resources. The act of co-creating marking criteria engages students in a deep understanding of knowledge (O'Donovan, Price, & Rust, 2008), because they must go back to study, review and look for information to create the guideline. Teachers can analyse the students' criteria and select the better descriptions when rewriting the final rubric.

2.- Engaging students in peer review. Authentic feedback processes improve students' ability to judge others´ work, as this is what is required in workplaces, thus developing evaluative judgment (Tai et al, 2017). Kearney & Perkins (2014) reported that 82% of their students considered seeing others' work in the process of peer marking promoted

better learning. In this context, peer review can be carried out when two peers collaboratively mark another student's anonymous test, judging their performance in the test. In these settings, a final grade may incorporate teacher assessment and students´ co-assessment (Tai et al, 2017).

3.- Using self-assessment in judging students' own work. Assessment can be more authentic when students are involved in dialogue and collaboration with their teachers in feedback processes (Bloxham & West, 2004). Kearney, Perkins, & Kennedy-Clark (2015) point out that in the first undergraduate year, students can self-assess, judge, mark and defend their own answers in a test in conversation with the teacher. Students develop an active role in constructing meaning with their teacher through an intersubjective relationship, exchanging and negotiating points of view (López-Pastor & Sicilia-Camacho, 2017; Lipnevich, Berg, & Smith, 2016). The following is an example of authentic feedback:

> In the class after the test, students define marking criteria in pairs. Each group presents to the class, and together they identify the main indicators and their description for three different levels of performance. Using those indicators, students review their own work and make qualitative comments about its strengths and weaknesses. Grades can be generated from weighting the teacher's evaluation of the test and the students´ comments.

Conclusion

Higher education must assess critical competences needed for solving realistic and contextualized problems using high-order skills in order that students become good

professionals and citizens. As tests are so widely used in higher education, this paper proposes changes to make them more authentic at three moments: planning, administering and follow up. While it may be desirable to lessen the overall weighting of tests in assessment regimes, and develop multiple forms of assessment, we have shown that some progress can be made towards designing tests that draw on the key dimensions of authentic assessment, and thus promote deep approaches to learning, more meaningful and engaging experience for students, and better preparation for the demands of work and life. Making assessment more authentic is a challenging process and will not occur without educational leadership and a desire to ensure that courses serve the needs of students beyond graduation.

Notes on contributors

Verónica Villarroel has PhD in Psychology. She works as a teacher and researcher at the Faculty of Psychology in Universidad del Desarrollo in Chile. She is the Director of The Center for Research and Improvement of Education (CIME), of the same Faculty.

David Boud has a PhD in Education. He has been involved in the research of higher education topics, for more than 30 years. He is an Emeritus Professor of the University of Technology, Sydney. He is also the founder, and Director, of the Centre for Research in Assessment and Digital Learning, of the University of Deakin in Melbourne, Australia.

Susan Bloxham has a PhD in Educational Research. She is currently an Emeritus Professor in the Faculty's Research Institute for Professional Learning in Education (RIPLE) following her retirement from the post of Director of the centre in the University of Cumbria, England.

Daniela Bruna has a PhD in Psychology. She works as a teacher and researcher of the Center for Research and Improvement of Education (CIME), which belongs to the Faculty of Psychology, of Universidad del Desarrollo in Chile.

Carola Bruna has a PhD in Biological Sciences. She works as a teacher and researcher of the Biochemistry and Molecular Biology Department, of the Faculty of Biological Sciences, in Universidad de Concepción, Chile.

References

Anderson, L., & Krathwohl, D. A. (2001). Taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of Educational Objectives. New York: Addison Wesley Longman.

Almond, R. J. (2009). Group assessment: Comparing group and individual undergraduate module marks. *Assessment & Evaluation in Higher Education, 34*, 141-148. doi: 10.1080/02602930801956083.

Altay, B. (2014). User-centered design through learner-centered instruction. *Teaching in Higher Education, 19*, 138-155. doi: 10.1080/13562517.2013.827646.

Ashford-Rowe, K., Herrington, J., & Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, *39*, 205-222, doi: 10.1080/02602938.2013.819566.

Avery, P. G., Freeman, C., & Carmichael, D. (2012). Developing Authentic Instruction in the Social Studies. *Journal of Research in Education*, *12*, 50-56.

Betts, L., Elder, T. J., Hartley, J., & Trueman, M. (2009). Does correction for guessing reduce students' performance on multiple-choice examinations? Yes? No? Sometimes? *Assessment & Evaluation in Higher Education, 34*, 1-15. doi: 10.1080/02602930701773091

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university: What the student does,* Maidenhead, Berkshire: Open University Press.

Bloxham, S., & West, A. (2004). Understanding the rules of the game: marking peer assessment as a medium for developing students´ conceptions of assessment.

Assessment & Evaluation in Higher Education, *14*, 20-30. doi: 10.1080/0260293042000227254.

Bosco, A. M., & Ferns, S. (2014). Embedding of authentic assessment in work- integrated learning curriculum. *Asia- Pacific Journal of Cooperative Education. 15*, 281-290.

Boud, D. (2010). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, *22*, 151-167. doi: 10.1080/713695728.

Brown G. A., Bull J., & Pendlebury M. (2013). *Assessing student learning in higher education*. Oxford: Routledge.

Carter, A., Sidebotham, M., Creedy, D., Fenwick, J., & Gamble, J. (2015). Strengthening partnership: the involvement of health care providers in the evaluation of authentic assessment within midwifery undergraduate education. *Nurse Education in Practice*, *15*, 327-32. doi: 10.1016/j.nepr.2015.01.013.

Davies, W. M. (2009). Groupwork as form of assessment: Common problems and recommended solutions. *Higher Education*, *58*, 563-584. doi: 10.1007/s10734-009-9216-y

Douglas, M., Wilson, J., & Ennis, S. (2012). Multiple-choice question test. A convenient, flexible and effective learning tool? A case study. *Innovations in Education and Teaching International, 49*, 111-121. doi: 10.1080/14703297.2012.677596.

Edstrom, K. (2008). Was doing course evaluations as if learning matters most. *Higher Education Research and Development, 27*, 95-106. doi: 10.1080/07294360701805234.

Endedijk, M. D., & Vermunt, J. D. (2013). Relations between student teachers' learning patterns and their concrete learning activities. *Studies in educational evaluation*, *39*, 56-65. doi: 10.1016/j.stueduc.2012.10.001.

Entwistle N. (2009). *Teaching for understanding at university. Deep approaches and distinctive ways of thinking*. Palgrave Macmillan: Basingstoke.

Flores, M. A., Veiga-Simão, M. A, Barros, A., & Pereira. D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: A study in Higher Education. *Studies in Higher Education, 40*, 1523- 1534. doi: 10.1080/03075079.20140881348.

Ghosh S., Bowles M., Ranmuthugala D., & Brooks B. (2017). Authentic assessment in seafarer education: Using literature review to investigate its validity and reliability

through rubrics. *WMU J Marit Aff, 15*, 317-336. doi: 10.1007/s13437-015-0094-0.

Gitanjali, M. (2016). The three Rs of written assessment: The JIPMER experience. *Journal of Pharmacology and Pharmacotherapeutics*, *7*, 115-119. doi: 10.4103/0976-500X.189650.

Guzzomi, A. Male, S., & Miller, K. (2015). Students´ responses to authentic assessment designed to develop commitment to performing at their best. *European Journal of Engineering Education*, *42*, 1-22. doi: 10.1080/03043797.2015.1121465.

Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *Curriculum Journal*, *16*, 207-223. doi: 10.1080/095851705001360093

Hart, C., Hammer, S., Collins, P., & Chardon, T. (2011). The real deal: Using authentic assessment to promote student engagement in the first and second years of a regional program. *Legal Education Review, 21*, 97-121.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77,* 81- 112. doi: 10.3102/003465430298487.

Hinton, D. P., & Higson, H. (2017). A large-scale examination of the effectiveness of anonymous marking in reducing group performance differences in higher education assessment. *PLoS ONE*, *12*, e0182711. doi: 10.1371/journal.pone.0182711.

Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: A test of student performance and attitudes. *Assessment & Evaluation in Higher Education, 37*, 125-136. doi: 10.1080/02602938.2010.515012.

Jensen, J., McDaniel, M., Woodard, S., & Kummer, T. (2014). Teaching to the test… or testing to teach: exams requiring higher order thinking skills encourage greater conceptual understanding. *Educ. Psychol. Rev*, *26,* 306-329. doi: 10.1007/s10648-013-9248-9

Johnson, D. W., & Johnson, F. P. (2009). *Joining together. Group Theory and group skills* (10th ed.). New Jersey, NJ: Pearson.

Kearney, S., & Perkins, T. (2014). Engaging students though assessment: The success and limitations of the ASPAL (Authentic Self and Peer-assessment For Learning) Model. *Journal of University Teaching and Learning Practice*, *11*, 1-13.

Kearney, S., Perkins, T., & Kennedy-Clark, S. (2015). Using self- and peer-assessments for summative purposes: analysing the relative validity of the AASL (Authentic

Assessment for Sustainable Learning) model. *Assessment & Evaluation in Higher Education*, *41*, 1-14. doi: 10.1080/02602938.2015.1039484

Lesage, E., Valcke, M., & Sabbe, E. (2013). Scoring methods for multiple choice ssessment in Higher Education—Is it still a matter of number right scoring or negative marking*? Studies in Educational Evaluation*, *39*, 118-193. doi: 10.1016/j.stueduc.2013.07.001

Ling Lau, K. (2013). Chinese language teachers' perception and implementation of self-regulated leaning-based instruction. *Teacher and Teaching Education*, *31*, 56-66. doi: 10.1016/j.tate.2012.12.001

Lipnevich, A. A., Berg, D. A., & Smith, J. K. (2016). *Handbook of human and social conditions in assessment*. Spain: Routledge.

López-Pastor, V., & Sicilia- Camacho, A. (2017). Formative and shared assessment in higher education. Lessons learned and challenges for the future. *Assessment & Evaluation in Higher Education*, *42*, 77-97. doi: 10.1080/02602938.2015.1083535.

Mahmoud F. A. (2014). A cross-cultural study of students' perceptions of assessment practices in higher education. *Education, Business and Society: Contemporary Middle Eastern Issues*, *7*, 293-315, doi: 10.1108/EBS-08-2014-0044.

Martinez-Rizo, F., & Mercado, A. (2015). Estudios sobre prácticas de evaluación en el aula: revisión de la literatura. *Revista Electrónica de Investigación Educativa*, *17*, 17-32. Download from: http://redie.uabc.mx/vol17no1/contenido-mtnzrizo-mercado.htm.

Martinez, M., O´Brien, M., Roberts, K., & Whyte, D. (2018). Critical pedagogy an assessment in higher education: The ideal of "authenticity" in learning. *Active Learning in Higher Education, 19*, 9-21. doi: 10.1177/1469787417723244.

Maxwell, T. (2012). Assessment in higher education in the professions: Action research as an authentic assessment task. *Teaching In Higher Education*, *17*, 686-696. doi: 10.1080/13562517.2012.725220.

McCabe, A., & O'Connor, U. (2014). Student-centred learning: The role and responsibility of the lecturer. *Teaching in Higher Education*, *19*, 350-359. doi: 10.1080/13562517.2013.860111.

McColongue, T. (2012). But is it fair? Developing students´ understanding of grading complex written work through peer assessment. *Assessment & Evaluation in Higher Education*, *37*, 113-123. doi: 10.1080/02602938.2010.515010.

Medland, E. (2016). Assessment in higher education: Drivers, barriers and directions for change in the UK. *Assessment & Evaluation in Higher Education*, *41*, 81-96. doi: 10.1080/02602938.2014.982072

Neely, P., & Tucker, J. (2012). Using business simulations as authentic assessment tools. *American Journal of Business Education*, *5*, 449-456.

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education, 39*, 102-122. doi: 10.1080/02602938.2013.795518.

O'donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: A nested hierarchy of approaches. *Teaching in Higher Education*, *13*, 205-217. doi: 10.1080/13562510801923344

O'Moore, L. M., & Baldock, T. E. (2007). Peer assessment learning sessions (PALS): An innovative feedback technique for large engineering classes. *European Journal of Engineering Education*, *32*, 43-55. doi: 10.1080/03043790601055576.

Partnership for 21st Century Skills. 2010. *American Management Association Critical Skills Survey*. Survey, Tucson: P21.

Pereira, D. Flores, M. A., & Barros, A. (2017). Perceptions of Portuguese undergraduate students about assessment: A study in five public universities. *Educational Studies*, *43*, 442-463. doi: 10.1080/03055698.2017.1293505

Rawson, K., Dunlosky, J., & Sciartellli, S. (2013). The power of successive relearning: Improving performance on course exams and long term retention. *Educational Psychology Review*, *25*, 523-548. doi: 10-1007/s10648-013-9240-4.

Raymond, J., Homer, C., Smith, R., & Gray, J. (2012). Learning through authentic assessment. An evaluation of a new development in the undergraduate midwifery curriculum. *Nurse Educational and Practice*, *13*, 471-476. doi: 10.1016/j.nepr.2012.10.006.

Richardson J. T. (2015). Coursework versus examinations in end-of-module assessment: A literature review. *Assessment & Evaluation in Higher Education*, *40*, 439-55. doi: 10.1080/02602938.2014.919628.

Schell, J., & Porter, J. (2018). Applying the science of learning to classroom teaching: The critical importance of aligning learning with testing. *Journal of food science education*, *17*, 36-41, doi: 10.1111/1541-4329.12141.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2017). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education, 76,* 467- 481. doi: 10.1007/s10734-017-0220-3

Tait-McCutcheon, S., Drake, M., & Sherley, B. (2011). From direct instruction to active construction: Teaching and learning basic facts. *Mathematics Education Research Journal, 23*, 321- 345. doi: 10.1007/s13394-011-0018-z.

Thornburn, M. (2008). Articulating a Merleau-Pontain phenomenology of physical education: The quest for active student engagement and authentic assessment in high-stakes examination awards. *European Physical Education Review*, *4*, 263-280. doi: 10.1177/1356336X08090709.

Vanaki, Z., & Memarian, R. (2009). Professional ethics: Beyond the clinical competency. *Journal Professional Nursing*, *25*, 285-291. doi: 10.1016/j.profnurs.2009.01.009.

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, *43*, 840-854, 2018. doi: 10.1080/02602938.2017.1412396.

Vu, T. & Dall´Alba, G. (2014). Authentic assessment for student learning: An ontological conceptualization, *Educational Philosophy and Theory*, *46*, 778-791. doi: 10.1080/00131857.2013.795110.

Wass, R., Harland, T., & Mercer, A. (2011). Scaffolding critical thinking in the Zone of Proximal Development. *Higher Education Research & Development*, *30*, 317-328. doi: 10.1080/07294360.2010.489237.

Watkins, D., Dahlin, B., & Ekholm, M. (2005). Awareness of backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lectures. *Instructional Science*, *33*, 283-309. doi: 10.1007/s11251-005-3002-8.

Wiggins, G., & McTighe, J. (2006). Examining the teaching life. *Educational Leadership*, *63,* 26-29.

Zhang, P., Ding, L., & Mazur, E. (2017). Peer instruction in introductory physics: A method to bring about positive changes in students' attitudes and beliefs. *Physical Review Physics Education Research*, *13*, 010104-1- 010104-9. doi: 10.1103/PhysRevPhysEducRes.13.010104.