Royal College of Surgeons in Ireland

repository@rcsi.com

Summative and Formative Assessment in Medicine: The Experience of an Anaesthesia Trainee

AUTHOR(S)

Sinead M. O'Shaughnessy, Pauline Joyce

CITATION

O'Shaughnessy, Sinead M.; Joyce, Pauline (2015): Summative and Formative Assessment in Medicine: The Experience of an Anaesthesia Trainee. Royal College of Surgeons in Ireland. Journal contribution. https://hdl.handle.net/10779/rcsi.10796819.v1

HANDLE

10779/rcsi.10796819.v1

LICENCE

CC BY-NC-SA 4.0

This work is made available under the above open licence by RCSI and has been printed from https://repository.rcsi.com. For more information please contact repository@rcsi.com

URL

https://repository.rcsi.com/articles/journal_contribution/Summative_and_Formative_Assessment_in_Medicine_The_Experience_of_an_Anaesthesia_Trainee/10796819/1

Summative and Formative Assessment in Medicine: The Experience of an Anaesthesia Trainee

Sinead M O'Shaughnessy[1] & Pauline Joyce[2]

[1] Specialist Anaesthesia Trainee Year 2, Department of Anaesthesia, Mercy University Hosptial, Cork

[2] Senior Lecturer, Royal College of Surgeons, Ireland

Correspondence: Sinead M O'Shaughnessy, Department of Anaesthesia, Mercy University Hospital, Grenville Place, Cork. E-mail: oshaugs@tcd.ie

Received: March 18, 2015 Accepted: April 8, 2015 Online Published: April 11, 2015

doi:10.5430/ijhe.v4n2p198 URL: http://dx.doi.org/10.5430/ijhe.v4n2p198

Abstract

The practice of medicine is becoming increasingly onerous for doctors, with growing demands coming from both patients and peers. Therefore, we need to equip doctors of the future with the appropriate skills to negotiate a very complex environment. Assessment should be central to this goal with a renewed focus on a more rounded examination process. Rather than a compete reliance on summative assessment, an increased focus should be placed on formative assessment, which has been historically neglected in medicine. In addition, it is important for educators to realise that a mutimodal approach is required to establish competence in a valid fashion and that each assessment modality should be used for both summative and formative assessment where possible. This will not only serve to address the external pressures placed on the teacher to streamline students but it will facilitate student learning as they become active participants in assessment.

Other areas of development in medical education going forward include peer assessment, which may be the answer to increasing cohort sizes. We also need to address the issue of maintaining adequate standards of assessment. Furthermore, we need to conduct further research to ensure that our assessment methods are an adequate predictor of future performance otherwise we as educational leaders will be failing both our students and the wider community.

Keywords: Medical education, Assessment in medicine, Assessment in anaesthesia, Anaesthesia trainees

1. Introduction

Assessment is such a pivotal aspect of education as it is a hugely powerful driver of learning for students. To the outside world, it represents the quality of students and institutions. Therefore, educators should strive to assess in the most beneficial way for students and to meet the changing demands of new generations and of the different assessment stakeholders. In this article, the writer would like to critically discuss assessment firstly giving a general overview of the topic, with particular focus on the area of anaesthesia. This paper will then look at both assessment of learning (AoL) and assessment for learning (AfL), discussing the various benefits and limitations of both approaches with particular reference to how these techniques are evolving. Furthermore, the challenges of assessment will be explored. Finally, a conclusion from the analysis will be presented focusing on the pertinent learning points in relation to the future of medical education.

2. Overview of Assessment

Assessment is defined as a process of collecting and evaluating information to measure students' progress (Brown, Bull & Pendlebury,1997, Bloxham & Boyd, 2007). It is well established that assessment shapes the experience of students and influences their behaviour more than any other element of their education (Snyder, 1971, Bloxham & Boyd, 2007, Epstein, 2007). It is a powerful tool in the armory of the educator and therefore, deserves careful consideration. When exploring assessment, there are six key questions which should be addressed; why, what, how, when, where and who (Harden & Laidlaw, 2012).

There are a number of reasons why educators need to assess. Firstly, educators are required to provide a means of discriminating between students. Doctors, particularly in the acute area of anaesthesia, must meet minimum safety standards. Assessment is also needed to provide institutional quality assurance to key stakeholders. Both of these goals are essentially AoL or summative assessment (SA) and have an outward focus serving wider societal aims

rather than the student. Placing the student in the centre of the examination process, however, we realise that assessment should foster and motivate student learning and equip them with the necessary skills to develop lifelong learning (Epstein, 2007, Bennett, 2011, O'Neill, 2014). This is AfL or formative assessment (FA) which serves to promote rather than simply measure learning. Each of these aims are equally important but can come into conflict. Ideally, we should have a rounded assessment process to address AoL and AfL simultaneously.

As educators, what blueprint do we use to assess our students so that they acquire the necessary knowledge, skills and attitudes? Constructive alignment ensures that the learning objectives are mapped carefully against assessment to ensure maximum possible validity (Hamdy, 2006). The four main elements, which need to be aligned, are content, intended learning outcomes, pedagogies and assessment principles (Biggs, 1999). When an assessment method lacks a comprehensive blueprint, there are a number of deleterious consequences for both educators and students. Construct under-representation is a pertinent example of this, which occurs when a topic forming only a small part of the curriculum is examined prominently. This not only undermines validity but also can cause huge distress to students (Messick, 1989, Coderre, Woloschuk & Mclaughlin, 2009). In order to avoid this, there are many models we can use for an assessment blueprint including Miller's pyramid (1990), Bloom's taxonomy (1956) and the Dreyfus and Dreyfus (1980) spectrum of skills acquisition to name but a few. The common thread to each of these is the evolution of student knowledge from the basic information of a novice to the richness of information of an expert (Bransford & Donovan, 2005, Schuwirth & van der Vleuten, 2011, Khan & Ramachandran, 2012).

The assessment methods commonly used in both undergraduate and postgraduate medical education are multiple choice questions (MCQ), extended matching questions (EMQ), essay questions, objective structured clinical examinations (OSCE) and oral assessment. These are used primarily as SA with a high stakes outcome. Each assessment method has a different strength in testing a student's knowledge, skills or attitudes but we need a carefully balanced combination to comprehensively reflect the assessment blueprint (Epstein, 2007, Table 1). Furthermore, the same modality can be used for both AoL and AfL and we should be careful not to miss important opportunities to fully develop the potential of a given assessment (Harlen & James, 1997, Epstein, 2007, Harris et al., 2007). For example, the membership examinations in anaesthesia, usually undertaken after one year of clinical experience, consist of two parts; the first involving MCQs and the second involving OSCE stations and an oral examination. These are assessed entirely summatively, however, which fails to capitalise on the maximal learning potential. By incorporating formative assessment through provision of feedback and stimulation of reflection, trainee learning could be exponentially increased. In addition, we also need to look at the 'one size fits all' traditional approach in how we assess students. An inclusive approach where students and teachers are equal partners in choosing how they are assessed has recently been advocated (Francis, 2008, O'Neill, 2010).

Table 1. Methods of assessment in medicine

Method	Domain	Type of Use	Limitation	Strength
MCQ	Knowledge	Summative	-Difficult to write -Cues	-High reliability -Computer graded -Efficient
SAQs	Knowledge	Summative Some formative	Reliability dependant on training of graders	-No cueing -Assess problem solving
Essays	Knowledge	Summative Some formative	-Time consuming -Interrater reliability	-No cueing -Higher order thinking required
DOPs/MINI-CEX	Skills Attitude	Formative Some summative	-Time consuming -Selective behaviours	Feedback by experts
Oral	Knowledge Attitudes	Summative Some Formative	-Subjective -Time consuming -Training of examiners	Feedback by experts
Simulator	Skills Attitudes	Formative Some summative	-Expensive -Can be artificial	-Tailored to educational goals -Often reliable and credible
Peer	Attitudes	Formative	-Confidentiality -Annonymity -Student buy-in	-Correlates with future clinical performance
Self	Knowledge Skills Attitudes	Formative	-Training required	-Fosters reflection and learning
Portfolio	Knowledge Skills Attitudes	Formative	-Time consuming -Student selects best material	-Fosters reflection and learning

Modified from Epstein, 2007.

Where and when to assess are also important considerations. A move towards work placed assessment would improve authenticity and help to examine attitudes, which lies at the apex of Miller's pyramid. Attitudes are very difficult in the traditional examination hall (Flin, Patey, Glavin & Maran, 2010, Miller & Archer, 2010). Similarly, the philosophy around when to assess is changing, with a move away from solely relying on end of year SA towards continuous intervals of examination in order to provide feedback for both the students and teacher. The use of portfolios, consisting of reflective commentary, is one method which can be used for continuous assessment which can have both SA and FA functions (Challis. 1999, Klenowski, Askew & Carnell, 2006).

Who should assess anaesthesia trainees is another pertinent question. With the resource limitations that exist in the Irish healthcare setting, which is undoubtedly reflected internationally, we should capitalise on the under utilised area of peer assessment (PA). It is well established that students learn from receiving feedback from their peers, as well as from giving feedback to each other (Ballantyne, Huges & Mylonas, 2002, Falchikov, 2007, Gielen, Dochy & Onghena, 2011, Kearney, 2013). Peer assessment may even stimulate deeper learning than in traditional assessment

settings given the intrinsic link between peer assessment and self reflection (Boud, Cohan & Sampson, 1999, Mulder, Pearce & Baik, 2014). When introducing peer assessment, we should be aware of the need for annonymity, however, which is a central concern of students engaging in this process.

3. Assessment of learning

Assessment of learning or summative assessment judges the overall progression of students in a systematic fashion (Biggs, 1998, Knight, 2002, Taras, 2005, Harlen & James, 1997). It is a high stakes assessment with a final mark of achievement awarded describing the learning achieved against public criteria. SA is a fundamental form of assessment as it allows comparisons to be made between students and it facilitates institutional profiling (Harlen & James, 1997). Although SA is regarded as not being 'student centric' and is often viewed as embodying all the negative social aspects of assessment, it allows students to proceed appropriately within the education system (Torrance, 1993, Black & Wiliam, 1998, Taras, 2005).

Summative assessment is the dominant force in both undergraduate and postgraduate medical education. Two of the biggest milestones for any anaesthesia trainee are the final medical examinations before graduation from university and the membership examinations in anaesthesia, usually undertaken after one year of clincial experience in the specialty. A combination of assessment methods are used in both of these examinations including MCQs, EMQs, essay questions, OSCEs and oral examinations. Using one author's experience in anaesthesia, these exams succeed in fulfilling certain educational objectives. They are hugely motivating for students due to the importance attached to the final grade (Brookhart, 2001). In addition, they are crucial for implementing minimum standards for professional practice. AoL fails the student in a number of areas, however, as it does not aim to further student learning or enourage students to take responsibility for their own educational needs. It can also create tension between the student and teacher (Gipps, 1994). How students cope with the task of passing these examinations is often through a combination of rote learning and memorisation which is a very superficial approach. SA encourages students to passively accepted ideas without necessarily understanding underlying theory and core foundation principles (Biggs, 1998, Ramsden, 2003). Furthermore, students often seek out preparatory courses looking for hints about what will be examined. Unfortunately, cue seeking behaviour is encouraged by SA (Miller & Parlett, 1974).

The final medical and anaesthesia membership exams differ in their examination styles. The final medical exams are primarily a criterion referenced assessment, grading students against a pre-specified benchmark, where as the membership examinations in anaesthesia are a norm referenced assessment, placing trainees onto a bell curve. Although norm referenced assessment is unavoidable where limited places are available for progression, it can be a very stressful process for students. Norm referenced assessment does not use explicitly described criteria, hence, it pits students against each other which places a huge sense of competition on the learning process (Biggs, 1998). To counteract this, medical educators should be aware that criterion referenced assessment allows for better constructive allignment and maintenance of standards (Cordiner, 2011).

How students perform in SA is very influential on their educational progress. This is why ensuring appropriate robustness of assessment methods used is so important (Harlen & James, 1997, Knight, 2002). Validity refers to whether an examination is 'fit for purpose' (Cronbach & Meehl, 1955). Reliability, on the other hand, is whether the exam is reproducible despite variability of examiners or of the day of examination (Kathirgamanathan & Woods, 2011, Schuwirth & van der Vleuten, 2011). MCQs are extremely popular in medical education at both an undergraduate and postgraduate level. This is unsurprising as MCQs are easy to write, store, administer, provide a meaningful assessment of cognitive ability as well as being highly reliable (Haladyna, 1999, Epstein, 2007). Educators need to be aware, however, that poorly written MCQs, facilitate success based on the savviness or testwiseness of the student rather than their knowledge (Downing, 2002). This contributes to construct-irrelevant variance and represents a threat to validity. Both Case and Swanson (1998) and Haladyna and Rodriguez (2013) have produced comprehensive guides to MCQ design to increase validity and reliability. As educators of the future, we must strive to create the fairest MCQs possible for our students particularly in light of their prominent use in high stakes examination.

To assess derives from the Latin verb 'assideo, assidere' meaning 'to sit beside'. Rather than sitting beside, SA conjures up images of the assessor passing judgement by standing over the student. In contrast, FA with it's predominant student focus fits this image better. I will now critically discuss AfL, which is starting to take a more central role in assessment of those practising within medicine.

4. Assessment for learning

Research from as far back as thirty years ago has demonstrated that feedback and student engagement are extremely powerful tools in stimulating learning (Ramaprasad, 1983, Sadler, 1989). AfL or FA is a process used by educators to foster student engagement through the provision of feedback to adjust and improve ongoing teaching and student learning (Natriello, 1987, Black & Wiliam, 1998, Heritage, 2010). In contrast to SA, FA focuses on learning rather than assessment. Feedback has been shown to raise student achievement but the quality of the feedback is key (Shepard, 2009). Appropriate feedback should not only identify the gap between the student's current and desired standards but should be actively used to close it by make changes in teaching practices and engaging the student (Archer, 2010).

Self assessment and PA are additional cornerstones of assessment for learning (Nicol & Macfarland-Dick, 2006). Self assessment has been described as a complementary feedback process which encourages students to monitor their own learning (Heritage, 2010). Self assessment in the form of both self monitoring and self regulation helps students to take responsibility for their learning and develop as self reliant lifelong learners. PA is mutally beneficial for both the assessor and the student receiving it (Wiliam & Thompson, 2006). By acting as assessor, a process that involves reflection around learning itself, it can further develop a student's understanding of their own learning. In addition, by what a student says about a colleague's work, we can easily decipher how well they have grasped the underlying learning goals and assessment criteria.

FA is under utilised in medicine, unfortunately to the detriment of students and trainees. As educators, we are now meeting a new cohort of students, generation Y, who crave timely feedback and are likely to reject SA as the key modality to fulfill their educational needs. In their important paper in 2001, Kathirgamanathan and Woods state that the purposes of assessment in anaesthesia are to ensure competence and fitness to practice. No reference is made to AfL and this is very much reflected in clinical practice. At the end of each six months rotation in anaesthesia, there is a formative assessment of trainee progress. The temptation by tutors, however, is to treat this as a box ticking exercise given the time and resource pressures faced within the clinical arena. Furthermore, it is far from ideal to allow six months to pass before critiquing trainees as effective feedback should be both timely and specific. All of this may be a reflection of the fact that many in the field of medicine lack qualifications in education and are thrust into positions as trainors without the appropriate understanding of their role (Ramsde, 2003).

Simulation training is recognised as a useful tool for clinical and performance based assessment and is widely employed as a training technique in anaesthesia (Motola, Devine, Chung, Sullivan & Issenberg, 2013, Gaba, Howard, Fish, Smith & Sowb, 2001). At six monthly intervals during anaesthesia training, trainees participate in a mandatory simulation session, which is formatively assessed. During these sessions, trainees are paired up and required to participate in a scenario which authentically recreates an anaesthetic emergency. These sessions act as powerful learning tools to help bridge the gaps in trainee knowledge particularly in relation to the non-technical skills of communication, teamwork and leadership. After each scenario, team feedback and debriefing occurs and arguably represents the most important aspect of simulation training (McGaghie, Issenberg, Petrusa & Scalese, 2010). Each participant along with their peers and the simulation trainers watch a recording of their performance and give feedback. Crucially, all events are linked back to the real life clinical environment. Simulation as an assessment method is so effective as the scenarios presented reside in the zone of proximal development of trainees. The complex resuscitation cases in simulation are often just beyond their comfort zones. By engaging the feedback with their trainers and peers, however, the students can further bridge the gaps in their knowledge, skills and attitudes to move towards greater clinical competence and independence.

5. Challenges of assessment

As medical practitioners, we are constantly required to assess the competence and performance of students, ourselves and our colleagues. The area of assessment in medicine is hugely challenging for clinicians and is likely to become significantly more demanding going forward. This is largely due to a combination of increasing cohort sizes and shrinking budgets (Fottrell, 2006, Gibbs, 2006).

When assessing students, certain aspects of competence are easier to assess than others. Knowledge can be reliably assessed using MCQs, EMQs and essay questions. Similarly, there are a number of effective frameworks for skills examination including OSCEs, directly observed procedures (DOP) and mini clinical evaluation exercises (MINI-CEX). In contrast attitude, comprising of teamwork, professionalism and communication skills, is more difficult to assess (Epstein, 2007). This presents a huge challenge given that quality of patient care is directly linked with a clinican's attitude (CQHCA, 2001).

Further challenges exist in relation to the inherent limitations of each individual assessment methods used in medicine. For example, MCQs are difficult to write and can result in cueing, essay questions can have high interrater reliability, oral examinations are subjective and may have sex and race bias and simulators are expensive and present artificial settings. To overcome these challenges, multimodal assessment can be used to provide broader insights into trainee competence and allows input from a number of assessors of the student's performance (Holmhoe, Sherbina, Long, Swing & Frank, 2010).

Standardisation of assessment is another issue to consider. Doctors are now more mobile than ever and medical migration is occuring worldwide, with greater movement towards more affluent countries (Bidwell at al., 2013, Chen & Boufford, 2005). How do we compare students or trainees from different medical education backgrounds when selecting for training posts? Of course, allowing individual medical schools to retain their autonomy regarding the examination process facilitates consistency between the curriculum and assessment (Colliver, Vu & Barrows, 1992). In the context of Irish medicine, however, huge challenges exist, particularly in relation to intern recruitment. Recent changes have created equal access for all students graduating from medical school within the EU for intern jobs in Ireland (Bidwell at al., 2013). However, it is unclear whether accreditation standards in relation to curriculum and assessment between the participating countries have been formulated. This is a real challenge to medical educators as it is well recognised that poor academic quality often results in deficient clinical care despite the underlying individual's level of experience (Hamilton & Pinnegar, 2000).

Finally, it remains very challenging to correlate assessment and future performance. There is a lack of robust evidence linking assessment with safe practice (Epstein, 2007). If this is what we expect of our future physicians, then we must strive to address this in order to produce competent doctors and ensure optimal patient care.

6. Conclusion

The practice of medicine is becoming increasingly onerous, with growing demands coming from both patients and peers. Therefore, we need to equip doctors of the future with the appropriate skills to negotiate a very complex environment. Asssement should be central to this goal with increased focus on FA, which has been historically neglected in medicine. In addition, it is important for educators to realise that the mutimodal approach is required to establish competence in a valid fashion and that each assessment modality should be used for both SA and FA where possible. This will not only serve to address the external pressures placed on the teacher to streamline students but it will facilitate student learning as they become active participants in assessment.

Other areas of development going forward include PA, which may be the answer to increasing cohort sizes. We also need to address the issue of maintaining adequate standards of assessment. Furthermore, we need to conduct further research to ensure that our assessment methods are an adequate predictor of future performance otherwise we as educational leaders will be failing both our students and the wider community. As stated by Harden and Laidlaw (2012), 'Students can walk away from bad teaching but are unable to do so with assessment if they are to achieve the qualification they seek'.

References:

Archer, J. C. (2010). State of the science in health professional education: effective feedback. *Medical Education*, *44*(1), 101-108. http://dx.doi.org/10.1111/j.1365-2923.2009.03546.x

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education*, *27*(5), 427-441. http://dx.doi.org/10.1080/0260293022000009302

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5-25. http://dx.doi.org/10.1080/0969594X.2010.513678

Bidwell, P., Humphries, N., Dicker, P., Thomas, S., Normand, C., & Brugha, R. (2013). The national and international implications of a decade of doctor migration in the Irish context. *Health policy*, *110*(1), 29-38. http://dx.doi.org/10.1016/j.healthpol.2012.10.002

Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: Granada Learning

Bloxham, S., & Boyd, P. (2007). *Developing Effective Assessment In Higher Education: A Practical Guide: A Practical Guide*. McGraw-Hill International.

Biggs, J. (1998). Assessment and Classroom Learning: a role for summative assessment? *Assessment in Education*, *5*(1), 103-110. http://dx.doi.org/10.1080/0969595980050106

Biggs, J.B. (1999) Teaching for Quality Learning at University. Buckingham: Open University Press.

Bloom, B. S., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain.

Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education*, *24*(4), 413-426. http://dx.doi.org/10.1080/0260293990240405

Bransford, J. D., & Donovan, M. S. (Eds.). (2004). *How Students Learn:: History in the Classroom*. National Academies Press.

Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education: Principles, Policy & Practice*, *8*(2), 153-169. http://dx.doi.org/10.1080/09695940123775

Brown, G., & Bull, J. M. Pendlebury (1997). *What is assessment? In assessing student learning in higher education*. London: Routeledge.

Case, S. M., & Swanson, D. B. (1998). *Constructing written test questions for the basic and clinical sciences* (2nd ed.). Philadelphia, PA: National Board of Medical Examiners.

Challis, M. (1999). AMEE MedicalEducation Guide No. 11 (revised): Portfolio-based learning and assessment in medical education. *Medical Teacher*, *21*(4), 370-386. http://dx.doi.org/10.1080/01421599979310

Chen, L. C., & Boufford, J. I. (2005). Fatal flows—doctors on the move. *New England Journal of Medicine*, *353*(17), 1850-1852. http://dx.doi.org/10.1056/NEJMe058188

Colliver, J. A., Vu, N. V., & Barrows, H. S. (1992). Screening test length for sequential testing with a standardized-patient examination: a receiver operating characteristic (ROC) analysis. *Academic Medicine*, *67*(9), 592-5. http://dx.doi.org/10.1097/00001888-199209000-00010

Committee on Quality of Health Care in America-CQHCA (2001). Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academy Press.

Coderre, S., Woloschuk, W., & Mclaughlin, K. (2009). Twelve tips for blueprinting. *Medical Teacher*, *31*(4), 322-324. http://dx.doi.org/10.1080/01421590802225770

Cordiner. (2011). Guidelines for good assessment practices. University of Tasmania. http://www.teaching-learning.utas.edu.au/__data/assets/pdf_file/0004/158674/GAG_v16_webversion.pdf. Accessed February 16th 2015.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281. http://dx.doi.org/10.1037/h0040957

Downing, S. M. (2002). Threats to the validity of locally developed multiple-choice tests in medical education: construct-irrelevant variance and construct underrepresentation. *Advances in Health Sciences Education*, *7*(3), 235-241. http://dx.doi.org/10.1023/A:1021112514626

Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition* (No. ORC-80-2). California Univ Berkeley Operations Research Center.

Epstein, R. M. (2007). Assessment in medical education. *New England Journal of Medicine*, *356*(4), 387-396. http://dx.doi.org/10.1056/NEJMra054784

Falchikov, N. (2007). The place of peers in learning and assessment. *Rethinking assessment in higher education: Learning for the longer term*, 128-143.

Flin, R., Patey, R., Glavin, R., & Maran, N. (2010). Anaesthetists' non-technical skills. *British Journal of Anaesthesia*, *105*(1), 38-44. http://dx.doi.org/10.1093/bja/aeq134

Fottrell. (2006). Report of the Working Group on Undergraduate Medical Education and Training. http://health.gov.ie/wp-content/uploads/2014/05/fottrell.pdf. Accessed 16th February 2015.

Francis, R. A. (2008). An investigation into the receptivity of undergraduate students to assessment empowerment, *Assessment & Evaluation in Higher Education,* 33:5,547—557. http://dx.doi.org/10.1080/02602930701698991

Gaba, D. M., Howard, S. K., Fish, K. J., Smith, B. E., & Sowb, Y. A. (2001). Simulation-based training in anesthesia crisis resource management (ACRM): a decade of experience. *Simulation & Gaming*, *32*(2), 175-193. http://dx.doi.org/10.1177/104687810103200206

Gibbs, G. (2006). Why assessment is changing. *Innovative assessment in Higher Education*, 11.

Gielen. S., Dochy. F., Onghena. P., et al,. (2011) Goals of peer assessment and their associated quality concepts. *Studies in Higher Education* 36(6), 719–35. http://dx.doi.org/10.1080/03075071003759037

Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. London: Falmer Press.

Haladyna. T. M. (1999). *Developing and validating multiple-choice testitems.* Lawrence Erlbaum, New Jersey.

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge.

Hamdy, H. (2006). Blueprinting for the assessment of health care professionals. *The Clinical Teacher*, *3*(3), 175-179. http://dx.doi.org/10.1111/j.1743-498X.2006.00101.x

Hamilton, M. L., & Pinnegar, S. (2000). On the threshold of a new century trustworthiness, integrity, and self-study in teacher education. *Journal of Teacher Education*, *51*(3), 234-240. http://dx.doi.org/10.1177/0022487100051003012

Harden, R. M., & Laidlaw, J. M. (2012). *Essential skills for a medical teacher: an introduction to teaching and learning in medicine*. Elsevier Health Sciences.

Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education*, *4*(3), 365-379. http://dx.doi.org/10.1080/0969594970040304

Harris, K-L., Krause, K., Gleeson, D., Peat, M., Taylor, C. & Garnett, R. (2007). *Enhancing Assessment in the Biological Sciences: ideas and resources for university educators*. Accessed 15 February 2015 from http://www.bioassess.edu.au/principles

Heritage, M. (2010). Formative Assessment and Next-Generation Assessment Systems: Are We Losing an Opportunity?. *Council of Chief State School Officers*.

Holmboe, E. S., Sherbino, J., Long, D. M., Swing, S. R., & Frank, J. R. (2010). The role of assessment in competency-based medical education. *Medical Teacher*, *32*(8), 676-682. http://dx.doi.org/10.3109/0142159X.2010.500704

Kathirgamanathan, A., & Woods, L. (2011). Educational tools in the assessment of trainees in anaesthesia. *Continuing Education in Anaesthesia, Critical Care & Pain*, *11*(4), 138-142. http://dx.doi.org/10.1093/bjaceaccp/mkr017

Khan, K., & Ramachandran, S. (2012). Conceptual framework for performance assessment: competency, competence and performance in the context of assessments in healthcare-Deciphering the terminology. *Medical Teacher*, *34*(11), 920-928. http://dx.doi.org/10.3109/0142159X.2012.722707

Kearney, S. (2013). Improving engagement: the use of 'Authentic self-and peer-assessment for learning'to enhance the student learning experience. *Assessment & Evaluation in Higher Education*, *38*(7), 875-891. http://dx.doi.org/10.1080/02602938.2012.751963

Klenowski, V., Askew, S., & Carnell, E. (2006). Portfolios for learning, assessment and professional development in higher education. *Assessment & Evaluation in Higher Education*, *31*(3), 267-286. http://dx.doi.org/10.1080/02602930500352816

Knight, P. T. (2002). Summative assessment in higher education: practices in disarray. *Studies in Higher Education*, *27*(3), 275-286. http://dx.doi.org/10.1080/03075070220000662

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2010). A critical review of simulation‐based medical education research: 2003–2009. *Medical Education*, *44*(1), 50-63. http://dx.doi.org/10.1111/j.1365-2923.2009.03547.x

Messick, S. (1989). Validity. In R.L. Linn (ed.), *Educational Measurement* (pp. 13–104). New York: American Council on Education and Macmillan.

Miller, A., & Archer, J. (2010). Impact of workplace based assessment on doctors' education and performance: a systematic rev. Giew. *Bmj*, *341*.

Miller, C. M., & Parlett, M. (1974). *Up to the Mark: A Study of the Examination Game.* Guildford: Society for Research into Higher Education.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic medicine*, *65*(9), S63-7. http://dx.doi.org/10.1097/00001888-199009000-00045

Motola, I., Devine, L. A., Chung, H. S., Sullivan, J. E., & Issenberg, S. B. (2013). Simulation in healthcare education:

a best evidence practical guide. AMEE Guide No. 82. *Medical Teacher*, *35*(10), e1511-e1530.

Mulder, R. A., Pearce, J. M., & Baik, C. (2014). Peer review in higher education: Student perceptions before and after participation. *Active Learning in Higher Education*, 1469787414527391. http://dx.doi.org/10.1177/1469787414527391

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, *22*(2), 155-175. http://dx.doi.org/10.1207/s15326985ep2202_4

Nicol, D. J., & Macfarlane‐Dick, D. (2006). Formative assessment and self‐regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199-218. http://dx.doi.org/10.1080/03075070600572090

O'Neill, G. (2010). Choice of Assessment Methods within a Module: Students' Experiences and Staff Recommendations for practice. *Methods*, *4*(42), 3-66.

O'Neill, G. (2014). Purposes & Principles of Assessment in Higher Education. University College Dublin. http://www.ucd.ie/teaching/resources/assessment/presentation/presentation.html. Accessed February 16th, 2015.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, *28*(1), 4-13. http://dx.doi.org/10.1002/bs.3830280103

Ramsden, P. (2003). *Learning to teach in higher education*. Routledge.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, *18*(2), 119-144. http://dx.doi.org/10.1007/BF00117714

Schuwirth, L. W., & van der Vleuten, C. P. (2011). General overview of the theories used in assessment: AMEE Guide No. 57. *Medical teacher*, *33*(10), 783-797. http://dx.doi.org/10.3109/0142159X.2011.611022

Shepard, L. A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, *28*(3), 32-37. http://dx.doi.org/10.1111/j.1745-3992.2009.00152.x

Snyder, B.R. (1971). *The hidden curriculum.* New York: Knopf.

Taras, M. (2005). Assessment–summative and formative–some theoretical reflections. *British Journal of Educational Studies*, *53*(4), 466-478. http://dx.doi.org/10.1111/j.1467-8527.2005.00307.x

Torrance, H. (1993). Formative assessment: some theoretical problems and empirical questions. *Cambridge Journal of Education*, *23*(3), 333-343. http://dx.doi.org/10.1080/0305764930230310

Wiliam, D. T., & Thompson, M. (2006). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), The future of assessment: Shaping, teaching and learning . Mahwah, NJ: Lawrence Erlbaum Associates.