A Preliminary Conceptual Model of the Assessment Design Process utilized by MCAST Science-Subject Lecturers

Richard Zahra
richard.zahra@mcast.edu.mt
Institute of Applied Sciences, MCAST

Abstract: The main purpose of this research study was the construction of a preliminary theoretical model that explores the interrelation of the knowledge, skills and beliefs that shape MCAST (Malta College of Arts, Science & Technology) lecturers’ approaches to assessment design. Educational research across many countries has demonstrated that assessment has a greater bearing on what students learn than teaching itself. Despite the large body of literature focusing on educational assessment, studies that attempt to explain the assessment design process through theoretical modelling are scanty. This research was intended to address that gap. A constructivist-driven grounded theory methodology was applied using multiple sources of data. Data was generated through recorded, semi-structured interviews with four MCAST science lecturers, selected purposefully from different academic backgrounds and with varying degrees of teaching experience. In addition, further data was collected from field observations and in-depth document analysis of assessment briefs. Through a process of concurrent data generation and analysis, a comprehensive analytical framework, grounded in the data, emerged. The assessment design process model resulted from the ‘actions/interactions and emotions’ domain of the analytical framework. It was subject to the ‘contextual conditions’ domain, which included the knowledge (assessment literacy), skills and beliefs of the participants. The latter directly influenced the ‘consequences/outcomes’ domain, the constructs of which were mostly derived from the analysis of assessment briefs which represented the end-product of the assessment design process itself. The findings of this study revealed that most of the participants’ approaches to assessment were rooted in the outmoded 20th-century dominant paradigm. These results concur with the findings of many other assessment-related research studies across the world that state that this paradigm still dominates the discourse in higher education assessment, despite the body of evidence pointing to more effective practices rooted in the espoused constructivist educational theories. These findings could thus be useful to formulate assessment improvement strategies which address the contextual conditions over which the organization has a degree of influence, and which in turn will have a cascade effect on the individuals’ assessment practices. The ideal outcome of such strategies would be the promotion of constructively aligned assessment practices and the prevalent adoption of ‘connected’ mental models amongst academic staff.

Keywords: Education, assessment, design, process, higher education, STEM, Constructive, alignment, literacy, mental, grounded theory, assessment design-process model, modes of assessment, vocational assessment, higher education assessment, assessment in science, STEM assessment, assessment tasks, MCAST assessment, assessment theory, theory in use, constructive alignment, constructivism, assessment paradigms, assessment for learning, assessment literacy, mental models, grounded theory.
Background to the Study

This study is an in-depth exploration of the knowledge, skills, and beliefs that shape MCAST science lecturers’ approaches to assessment. The main goal was the construction of a preliminary substantive theoretical model to explain how these cognitive subdomains interrelate during the creative design process of assessment tasks. Despite the multitude of studies focusing on educational assessment, the assessment design process itself has not yet been thoroughly explained through theoretical modelling. This study meant to address that gap in educational literature.

Assessment is a fundamental aspect of any educational organization. Quinn (Clarence et al. 2015) states that assessment has a greater bearing on what students learn than teaching itself. Students perceive assessment as an indicator as to what is really worth learning. Quinn adds that ‘Assessment is thus, for most students, a ‘lever’ which determines how and why they will learn in a course.’ Whilst lecturers see assessment as the last link in their chain of priorities (Curriculum Outcomes - Teaching Activities - Assessment), students perceive assessment at the head of their priority list (Assessment - Learning Activities - Curriculum Outcomes) (Biggs 2003). Assessment therefore frames how students react to their learning environments and often determines their actions. It is also an indicator of an institution’s priorities for making educational judgements and provides an agenda that is more persuasive than a unit specification or curriculum. As a consequence, assessment has a ‘backwash effect on all teaching and learning activities’ (Boud 2007), exerting a key role on the effectiveness of any educational organization.

MCAST is a vocational college established in 2001 and was the first Maltese educational institution to adopt an outcome-based educational model (OBE). It is the only higher education organization in Malta in which learning outcomes (LOs) and associated assessment criteria are central to the teaching, learning, and assessment processes. This approach follows the standards model of assessment (criterion-referenced assessment) that is considered more in tune with currently – espoused theories of learning than the measurement model of assessment, also known as norm-referenced assessment (Biggs 1995). Every assessment task is internally verified before being officially issued. Verification is usually carried out by other lecturers and provides an inbuilt quality assurance system that goes some way to ensuring that assignment tasks are aligned with LOs and assessment criteria. The MCAST situation is unique when considering that it has adopted an avant-garde educational assessment approach which is, however, implemented by a lecturing body whose majority do not hail from an educational background. Although this factor has certainly helped to create close ties with the very industries that employ MCAST graduates, one concern that arises is the aptitude and preparation of academic staff with regards to assessment literacy.

Since its first advocacy during the 1960s (Morcke et al. 2012), OBE has been internationally favoured and implemented in countries such as Canada, the United States, and New Zealand, and is regarded as an approach that favours educational renewal (Malan 2000). Locally, the education authorities have recently extended this approach to schooling at all levels, and a ‘Learning Outcomes Framework’ (LOF) is currently being introduced in all primary and secondary schools (Schoolslearningoutcomes.edu.mt 2018). William Spady, regarded as OBE’s leading exponent, defines outcomes as: ‘… clear learning results that we want students to demonstrate at the end of significant learning experiences … what learners can actually do with what they know and have learned – they are the tangible application of what has been learned’ (Spady 1994: 2) Thus, the early adoption of OBE by MCAST could be seen as way of preparing the students with a knowledge base and skill-
set that will enable them to transfer what they have learnt in the lecture rooms to their prospective work places.

The research question asked was: How do relevant knowledge, skills, and beliefs integrate together to influence the MCAST science lecturers’ assessment design process? Research has demonstrated that the approach to assessment depends on relevant knowledge, skills, and the designers’ beliefs, including their philosophical world-view and the learning theories they adhere to explicitly or implicitly. All these factors enable assessors to enact mental models that guide their approaches to assessment (Heinrich 2017). The study involved a qualitative investigation of the participants’ level of pedagogical content knowledge, their possession of practical and cognitive skills related to effective assessment design and implementation, and an examination of the philosophical and pedagogical beliefs, if any, that guided the participants’ approaches to assessment. Thus their mental models were deconstructed and analysed. Through the constant comparative analysis of generated empirical data, a preliminary theoretical model of the assessment design process was constructed to shed light on what measures the individual assessors, as well as the organization as a whole, could adopt to enhance the quality of assessment approaches. Such a theoretical model could also highlight the requirements of an effective induction experience in assessment practices for newly employed lecturers. In addition, the findings of this study could provide valuable insight as to the challenges the Maltese educational system will be facing in the coming years when introducing innovative forms of assessment that are a necessary prerequisite for the successful implementation of the ‘Learning Outcomes Framework’ (LOF) at all levels (Schoolslearningoutcomes.edu.mt 2018).

**Literature Review**

Comprehensive knowledge of a discipline does not imply the ability to teach and assess its constructs effectively. In addition to content knowledge, the three areas necessary for basic teacher-training include: pedagogical content knowledge, learning theory, and assessment literacy (James 2006). All three areas are interrelated. For instance, the choice of assessment approaches a lecturer takes will be based on the learning theories that the individual adheres to, often implicitly, and which is also reflected in the pedagogy used. Such choices are crucial in determining the effectiveness or otherwise of the assessment devised by educators. Heinrich (2017) affirms that, in higher education systems, learning and assessment require educators to adapt to multiple influences and to the pressures of the learning environment. All these influences will consequently lead educators to enact mental models that will ultimately guide their choices. Rook (2013) describes a mental model as ‘a concentrated, personally constructed, internal conception of external phenomena (historical, existing, or projected) or experience, that affects how a person acts’. Educators frequently construct instruction and assessment without fully understanding the ways their own experiences, training and current contexts influence their practices (Heinrich 2017). Exploring the mental models of individuals in an organization is a step forward towards institutional improvement, which can lead to changes in the behaviour of individual members and their practices (Rook 2013). According to Heinrich (2017), three categories of influence bear on lecturer’s assessment practices. These are training and disciplinary socialization (the accepted standards for a particular discipline), environmental and cultural socialization (the less formal influences of the work environment and the group), and incentives and accountability (the more formal institutional accountability demands and accreditation measures).
Biggs (1995) stipulates three dimensions that interact to yield different modes of assessment. These are 1) the function of testing – whether a Measurement or the Standards model of assessment is adhered to; 2) Quantitative vs. Qualitative assumptions on the nature of what is to be learned and assessed; 3) whether learning is situated or decontextualized. The Measurement model of assessment, also known as norm-referenced assessment, involves ranking students' achievement in order to compare student with student. The rank placement will determine the grade achieved, denying the effect that good teaching and learning can make on student achievement. By contrast, the Standards model (e.g. criterion-referenced assessment) involves assessing the achievement of set standards (learning outcomes and assessment criteria). The underlying assumptions of the Standards model are that public standards can be set which can be reached by most students but at different levels of performance and that fair and consistent judgements are possible on whether these standards have been met or not. The Standards model is the basis of assessment in outcome-based organizations like MCAST. The Quantitative tradition of assessment stems from the positivist tradition in social sciences (Moss 1992) and conceives learning as the acquisition of 'specific discrete skills, described as well-delimited behaviours' (Cole 1990), that is quanta of knowledge. In assessment terms, achievement is the measurement of how many discrete units (facts, skills, competencies, behavioural objectives, and performance indicators) have been accumulated by a learner. This view is based on Behaviourist learning theories. On the other hand, the Qualitative tradition of assessment is based on the Constructivist family of learning theories. Learners are construed as active, seeking meaning by constructing knowledge and adapting existing conceptual structures, rather than by receiving and storing knowledge. The teacher's task in such a framework is not be to transmit correct understandings but to help students construct progressive understandings that are conceptually more developed and congruent with accepted thinking.

These conceptual frameworks, constructed through personal experiences, will help students make sense of the world around them (Biggs and Moore 1993). Therefore, assessment in such a context implies charting the progress of these individual conceptual models over time, from a basic to a high level of competence. For Biggs (1995) Qualitative assessment can take two forms: 1) developmental assessment to discover where students stand in the development of understanding or competence in the concept or skill in question – the focus is on discipline-based or declarative knowledge; 2) ecological assessment that assesses the student's ability to use knowledge in practical terms – the focus is on applications and problem-solving. The latter requires situated assessment tasks that assess the learners' ability to apply procedural knowledge in an authentic context. This is a preferred mode of assessment since it requires students to think, take decisions, and pursue courses of action that are reflective of real-world situations (Archbald and Newmann 1988). However, there is still a place for decontextualized learning and assessment as a significant amount of legitimate declarative (theoretical) knowledge is taught at all levels of the educational system, and this still requires assessing. Biggs (1996) and Heinrich (2017) both cite Argyris's (1976) exploration of the relationship between what an individual says is important and how that individual actually acts in the organization. Espoused theories are those accepted by the educational community, which underlie professional practice. However, it is theories-in-use that actually guide the practice of individuals. When the theory-in-use of an individual does not match espoused theory, this has a significant bearing on the quality of the individual's output and, consequently, on the effectiveness of the organization as a whole. This disjunction between espoused theory and theory-in-use is common amongst educators, as is lack of theoretical coherence (James 2006). Espoused educational theories are broad, encompassing not only theories of teaching and learning, but also theories of the nature of learning (Biggs 1996).
Shepard (2000) stipulates that all assessment practices are rooted in one of two major paradigms. The twentieth century dominant paradigm was characterized by: 1) a Social Efficiency Curriculum (that assumed that the principles of scientific management that were intended to maximize efficiency of factories could be applied with equal success to educational institutions); 2) Principles of Scientific Measurement (involving the application of scientific measures to maximize efficiency in education by, for example, predicting who is best suited for each job and grouping learners according to ability); 3) Behaviourist learning theories (which conceived of learning as mastery of bits of knowledge through an accumulation of stimulus-response associations, and which required frequent testing to ensure mastery). These theories led to the widespread adoption of 'objective testing' in classrooms throughout the twentieth century, which shaped many of today's educators' beliefs about the nature of evidence required and principles of fairness, particularly the notion that assessments have to be uniformly administered to all students.

Towards the latter decades of the twentieth century, there were significant advances in the understanding of how learning takes place through the work of researchers from different disciplines including linguistics, computer science, cognitive science, and neural science (James 2006). These discoveries led to an emergent paradigm of assessment and learning that started in the 1990s and is still accepted today as the espoused set of theories (Shepard 2000). The emergent paradigm is characterized by: 1) A reformed vision of the curriculum (that is in line with the Social-Constructivist conceptual framework); 2) Classroom assessment that promotes learning, the use of higher-order thinking skills and knowledge transfer, and acknowledges a social dimension to learning; 3) Constructivist learning theories (that imply a qualitative approach to assessment).

Research has shown that the twentieth-century paradigm of assessment was leading to adverse results in student learning and motivation. To cite one example, recent studies across the USA have shown that college graduates have retained little understanding from their science courses, and even the brightest students graduating from prestigious institutions like Harvard were found to foster serious misconceptions regarding basic scientific concepts (Lord and Baviksar 2007). Yet, despite the conclusions from research, the dominant discourse of assessment in higher education institutions is measurement and certification, which construes learners as passive subjects (Boud 2007). This dominant discourse is rooted in the outdated twentieth-century dominant paradigm of assessment and the assessment practices of many practitioners are still stuck in the 1970s. This is particularly evident in traditional lectures, where the associated student activities are listening, interpreting, comprehending, note-taking, and, sometimes, reflecting. The problem with traditional lecturing is that it usually offers a narrow range of learning-related activities, so instances of high-level engagement are either the result of individual student brilliance or of chance (Biggs 1996). Biggs's (2003) remedy to this situation is constructive alignment, wherein all components in a teaching system – the curriculum and its intended outcomes, teaching methods, and assessment tasks – are aligned seamlessly to each other. If both the curriculum and teaching/assessment practices are aligned to Constructivist theory and to each other, the learner will be ‘trapped’ into learning what is intended to be learned. In such a framework, the lecturer’s role will not be that of knowledge imparter, but a catalyst for learning. Ultimately, ‘... what the student does is actually more important in determining what is learned that what the teacher does' (Shuell 1986, cited in Biggs 2003).
The Behaviourist paradigm supports summative assessment of learning as its preferred mode of evidence gathering. On the other hand, the Constructivist paradigm entails assessment for learning (AFL), that is a formative type of assessment in which the ultimate goal is not just the description of student learning, but also the enhancement of that very learning. In higher education, however, the emphasis still remains on assessment of learning rather than AFL (Quinn in Clarence et al. (eds.) 2015). In such a scenario, students remain passive recipients of the assessment acts of their tutors (Boud 2007). However, when they move on to work contexts, that is contexts for which higher education is intended to prepare them, they necessarily have to reinvent themselves as active subjects and learn to self-assess and also interact with their peers for improvement. Assessment is truly ‘formative’ when the evidence generated is actually used to adapt the teaching approach to meet the learning needs of the students (Black et al. 2004). In 1998 Black and Wiliam reviewed evidence published in over 250 articles by researchers across several countries and found a strong body of evidence claiming that formative assessment raises standards significantly.

Later, in 2004, Black et al. carried out intervention experiments to measure the kinds of gains that could be achieved in classrooms using formative assessment for extended periods of time. After training teachers in a number of schools in practical AFL techniques, they compared the performances of classes exposed to AFL to others that were served with the traditional assessment of learning. They discovered that the average improvement effect size was 0.3 standard deviations for the treated classes. Yet, despite the strong evidence pointing to the learning and achievement benefits of AFL, including the fostering of critical thinking and problem-solving skills, which are so much in demand by employers worldwide, most educational practitioners still use assessment practices rooted in the twentieth-century dominant paradigm. In higher education, for instance, assessment is still predominantly related to issues of measurement, certification, and quality assurance (Boud 2007), rather than used to promote learning. On the other hand, effective AFL strategies can go a long way to creating a lecture atmosphere wherein higher-order thinking is evidenced and shared by all students. In conclusion, effective assessment is a contextualized one. Wiggins (1992) states that typical assessment tasks, even those considered demanding, tend to over-assess student knowledge and under-assess higher-order skills and rarely put students in an authentic performance situation, where thinking, not just obvious bits of knowledge, is required. He suggests that tasks should be contextualized through authentic simulations of real-life problems whose solutions support multiple approaches and require good judgements in achieving effective results.

Research Methodology

A qualitative (grounded theory) methodology was used in the implementation of this study. Data generation mainly took the form of recorded semi-structured interviews with four selected science MCAST lecturers. Purposive sampling (Birks and Mills 2015: 11) was used to select participants with varying degrees of experience in teaching. Additional sources of data were also used, such as field observation and document analysis (each participant was asked to provide what he considered his best assignment brief which represented the outcome of the assignment design process being investigated). The researcher was in a ‘backyard research’ situation (Glesne and Peshkin 1992, cited in Creswell and Creswell 2018), that is a study within his own organization and immediate work setting. He was thus very familiar with the context. Throughout the study, essential grounded theory methods as suggested by Birks and Mills (2015: 10) and Chong and Yeo (2015: 262–3) were implemented to ensure ‘procedural logic’ (Birks and Mills 2015: ...
A process of concurrent data generation and analysis was implemented, wherein each interview was transcribed, coded, and analysed before the subsequent interview was carried out. Incoming data generated from subsequent interviews and document analysis was integrated in the existing analytical framework so that the process of coding and category development was progressively elevated to a higher level of abstraction (constant comparative analysis). The data was questioned during the analytical process using Corbin and Strauss's (2008) Coding Paradigm as a structural map for the conceptual framework. Initially, codes were grouped under 3 categories: 1) Contextual conditions; 2) Actions, interactions, and emotions; 3) Consequences. MAXQDA software was used to keep track of the coding and analytical processes, and the reflexive progress of the researcher was documented within the memoing system of the program. The final outcome was the synthesis of a preliminary conceptual model that attempted to explain the assessment design process. A constructivist approach to grounded theory was adopted since this epistemological stance places a focus on the relationship of the author with the participants and acknowledges the central role of the researcher and his experiences in the study (Charmanz 2014 cited in Birks and Mills, 2015: 7). The author affirms that Constructivism has influenced his vision for science education for the past 24 years. The tenets of Constructivism have informed his professional practice and choices of pedagogy and assessment methods. Thus, the interpretation of the findings were shaped by the researcher's own professional experience and academic background. Through an inductive conceptualization approach, an attempt was made to generate meaning and explain the process of assessment design from the data collected in the field (Creswell and Creswell 2018). The grounded theory approach satisfied three basic preconditions: 1) Not a lot of research has been done on the assessment design conceptual process; 2) the generation of theory with explanatory power could be useful in explaining this process; 3) the process itself, which is abstract in nature, was well suited to be investigated by grounded theory methods since these would result in a conceptual explanation derived from the actual data (i.e. the participants' perspective). The grounded theory approach to this investigation was a 'methodologically congruent' choice (Birks and Mills 2015: 36) since there was convergence between the author's philosophical position, the aims of the study, and the methodology used.

Measures were also taken to ensure that the study was ethically sound. The anonymity of the participants and their personal information were respected at all times. Since the data extracted from the interviews was fractured and reassembled into an abstract conceptual form during the study, any links to individual participants were extinguished. Precautions were also taken to ensure that electronically stored confidential information would not be disclosed to third parties (Aera.net 2011). Since social researchers are under the obligation of protecting subjects from undue harm from their participation in research, including moral harm (The-sra.org.uk 2003), measures were therefore taken to minimize the possibility of moral harm occurring to participants. Since unstructured interviews can lead to a degree of intrusion on the participants' professional and personal domains (The-sra.org.uk 2003), potential subjects were asked to participate through informed consent. An interview covering letter explaining the objectives of this research study, the methodology used to generate data, and a request for their permission to be interviewed was given to each potential participant. Participation was strictly voluntary, and the researcher collected a duly signed consent form before the start of each interview. Conditions related to confidentiality and anonymity were also explained verbally to each participant. The same interview protocol was used for all participants (Creswell and Creswell 2018) during the generation and collection of data. In conclusion, the benefits of the study for the participants as well as the institution were explained.
Presentation and Analysis of Emerging Constructs

Analysis of the data from interviews and assignment documents resulted in the construction of a basic conceptual model of the assignment design process. The analytical framework is depicted in Figure 1. The major code categories that emerged from the analysis (Figure 2) led to the formulation of the conceptual process model discussed in further detail in the subsequent section.

Figure 1: Analytical framework
This process model emerged in the *Actions, interactions, and emotions* domain of the coding paradigm, as it is the result of the participant's actions which are, in turn, influenced by their interactions with others and the institution, as well as their emotions (beliefs and values). The categories that emerged in the *Context/conditions* domain of the analysis pointed to the participants' assessment-related knowledge, skills, and beliefs that ultimately influenced various stages of the assessment design process. The *Consequences/outcomes* domain represented the end-product, that is the assessment brief, but not the end-result of the process since most participants took further action (Figure 1). Data for the development of codes and categories in the consequences/outcomes domain was mostly generated by the analysis of the assignment documents provided by each participant. The documents also served to verify and validate some of the participants' assertions.

**Figure 2:** Major domains and associated major code categories

**A preliminary assessment design process model**

Figure 3 represents the major categories (high-level concepts) that emerged through an inductive analysis of the data and that characterized the assessment design process. At the *conceptual planning stage* (see Figure 1 for depiction of sub-categories), the assessor envisages the approach he will take to create the assignment.

**Figure 3:** The assessment design process
Factors such as practical considerations and strategic and pedagogical planning come into play in determining the approach of the assessment designer. Practical considerations were ranked high by participants as they presumably helped them deal with the pressures and constraints that they habitually face, such as workload and time constraints. These practicalities included curriculum requirements (mostly learning outcomes and success criteria), articulating their own goals and working backwards from them, and taking into account the students’ perceived needs, abilities, or inclinations to learning. Thinking that involved strategic and/or pedagogical planning was used by participant lecturers to decide what types of tasks to include, the level of cognitive skill to set these tasks at, and whether to contextualize or not the assignment as a whole or some individual tasks. The approaches of the participants were often not implicitly declared and these were aligned either to Behaviourism or to Constructivism. The underlying philosophical views of the participants and hence their approaches were inferred from the conceptual analysis of their interviews as well as the assignment briefs. Most participants, as expected, adopted Behaviourism as their theory-in-use (Biggs 1996), though this was very implicit and never affirmed explicitly. One participant was in the process of actively adopting a Constructivist stance and was struggling to shed the Behaviourist beliefs that had been instilled in him through a long educational experience, during which he was mostly assessed along the lines of the twentieth-century dominant assessment paradigm. Complete constructive alignment, that is alignment of teaching activities and assessment activities to both Constructivist theory and to the learning outcomes was attempted by only one of the participants. On the other hand, all participants considered alignment of assessment tasks to assessment criteria as a matter of accountability (though none mentioned that such an alignment is required for the validity of a norm-referenced task). Creativity and the need to empathize with students were also notions that some participants thought important during the strategic planning of their assessments. Wiggins (1992) discussed the need for creativity in designing ‘performance assessment’ by calling upon assessment makers to be creative designers, not just technicians. Tasks set in authentic contexts that the students can associate with require a significant degree of creativity to produce.

The resource utilization stage involved the use of resources in the preparation of the assignment brief. Participants mentioned the use of books, websites, popular media (such as online newspapers), and academic journals to get ideas and contexts for their assignment tasks. Only one of the participants produced visuals in his assignment in order to make question contexts clearer and visually appealing for the students. All others used only text-based tasks. After collating resources, the participants embarked on the writing stage. During this phase, the designer produced the first drafts of the assignment. The amount of effort and tools used (such as ICT) depended on the individual’s preferences and motivation. A minority of the analysed tasks were paraphrases of the assessment criteria. One participant wanted to produce visually appealing task sheets, while another included detailed expectations and guidelines in his assignment brief. The reviewing stage entailed re-reading first drafts and checking for errors and inconsistencies. Most participants carried out this exercise before handing in the brief to the internal verifier for reviewing. Some participants trial-ran the assignment by working it out and some even produced marking schemes or rubrics at this stage. Others checked the face validity of the tasks by having an experienced peer (a mentor figure, not the verifier) go through their assignment. The participants all checked that the tasks were aligned to the assessment criteria, an important institutional accountability factor at MCAST. During the implementation stage, the assessor put into motion the assessment. The form of assessment (research-based, test, or practical-based) affected the implementation process of the participants and required knowledge of particular procedures. An important consideration here was that the assessor might have to overcome stumbling blocks that he unexpectedly encountered during the implementation of the assessment.
like, for instance, logistical considerations for a test involving more than one cohort of students.

After correcting each assignment, assessors would then embark on the *feedback giving/getting stage*. Feedback is arguably the most important instructive action that affects future learning and achievement (Rust *et al.* 2005). Feedback was mostly given in the form of marks and sometimes, participants also went through the corrected assignment with the students. Feedback in the form of marks or grades, however, has been proven to be of limited effectiveness and sometimes even counterproductive to student learning (Butler 1987). Also, if feedback is given too long after the work was done, the students lose interest and they are unlikely to act upon it. When the feedback emphasis is on the mark, Wotjas (1998) has shown that it may be perceived to relate to the student's personal worth as a person, rather than to the individual piece of work. According to Sadler (1989), the conditions for effective feedback include: 1) Knowledge of standards; 2) comparing those standards to one's own work; 3) taking action to close the gap between the two. Feedback should also be a two-way process and some participants mentioned that student assessment tasks provide feedback to them about their own teaching and assessment methods. If the assessor considers himself as a reflective practitioner, he may choose to reflect upon the effectiveness or otherwise of the assessment process – the *reflection stage*. Professionalism entails reflective practice but in the MCAST context (as in most levels of the Maltese educational system) there are no widespread accountability incentives that promote this practice. Environmental socialization (Heinrich 2017), however, has a bearing on whether assessors are encouraged to be reflective practitioners or not through the influence of their peers. Reflective practice is metacognition at work – the process involved when individuals plan, monitor, evaluate, and make changes to their own learning behaviours (Cambridge assessment 2018). Assessors who use best practices could be classed as *reflective learners* (Perkins 1992) who use strategic thinking, reflect upon learning, and evaluate the strategies they are using and revise them as appropriate. After reflecting, the assessor has then to decide whether or not to act upon the feedback received (purposeful or otherwise). If so, he will have reached the *reaction stage* of the assessment process.

The analysis of the participants' interviews revealed that this involved measures by the lecturer to change tack regarding either the assessment approaches used, the classroom pedagogy that led to the assessment, or both. The reaction to summative assessment, however, has no bearing on the assignment administered, although it might influence future assessments and teaching approaches. By contrast, the timely reaction to feedback gained through *assessment for learning* methods can be effectively used to improve the performance of the students in the very same assessment process (Black *et al.* 2004). The beliefs and personal philosophy will influence how an assessor reacts to feedback or even his lack of reaction (to preserve the perceived comfort of the status quo). Interestingly, one of the participants highlighted the notion that embarking on the reaction measures is a way of making the assessment practices more credible in the eyes of the students and thus strengthens the ethos of the whole MCAST organization. In conclusion, it is worth noting that whereas the first 6 stages of the assessment design process are likely to be carried out, at least to some degree, by all lecturers due to formal accountability measures taken by the institution, including the internal verification process, the *reflection and reaction stages* are less likely to be implemented by all assessors. Their implementation or otherwise will depend on the enacted mental model of the individual, which in turn will be influenced informally by environmental and cultural socialization, such as the ongoing influences of peers in the same discipline, institute practices and expectations, and belief in the working towards the ‘greater good.’ (Heinrich 2017).
Influences on the Effectiveness of Assessment Design

The contextual conditions that emerged from the analysis had a direct bearing on the approaches that participants took in assessment design. The main concepts that emerged in the ‘context/conditions’ paradigm were categorized into assessment-related knowledge (assessment literacy), skills, and beliefs (Figure 1). Assessment literacy is described by James (2006) as one of the three pillars of basic teacher training, together with pedagogical content knowledge and learning theory. Mertler (2003) states that assessment literacy can be defined as ‘the possession of knowledge about basic principles of assessment and evaluation practice, the development and use of assessment methodologies and techniques in the classroom, familiarity with different tools and apparatus of assessment, familiarity with standards of quality in classroom assessment, and familiarity with an alternative to traditional [Behaviourist-based] measurements of learning’ (Mellati and Khademi 2018). Since all four participants did not have educational backgrounds, the only formal training they had been exposed to was through the Vocational Teacher Training Unit (VTTU) which is a compulsory training module for MCAST lecturers that determines salary scale progression. The level of pedagogical content knowledge amongst participants was low even amongst the most experienced, as none had been trained in the methodology of science-subject teaching (the teaching experience of the participants ranged from 1 to 15 years). Three out of the 4 participants had been involved in curriculum design (unit writing) at MCAST. However, the only participant who had implemented assessment and teaching methods stemming from espoused theory (Constructivism) was relatively inexperienced (4 years’ teaching career), but he had been mentored informally by a colleague with extensive practical experience of innovative teaching and assessment methods. By contrast, most participants demonstrated a degree of ability to interpret and apply Bloom’s taxonomy of educational objectives but most were less well-versed in the use of assessment tools that could be used in the context of the emerging paradigm of assessment (Shepard 2000). Assessment skills were often expected to be developed on the job.

The implicit beliefs of participants about assessment had a significant impact on their approaches to assessment. All participants had qualifications at masters’ level and beyond and the data strongly pointed out that the assessment methods they had been exposed to as students during their formative years had influenced and shaped their outlooks and often dictated their courses of action. This would explain the implicit but deeply rooted endorsement of Behaviourist learning theory (nowadays considered to have adverse impacts on student learning and achievement) that guided most of the participants. Changing the view radically to endorse Constructivist theory-in-use would be a very arduous step since it would involve changing long-established habits and moving into unfamiliar territory (Wiliam 2010). When specifically asked about their vision or philosophy of assessment, none of the participants related to any of the world views that interpret the nature of learning in academic research contexts, such as Post-Positivism or Constructivism (Creswell and Creswell 2018). Instead, they articulated their visions in a pragmatic manner by giving statements such as ‘based on my experiences as a [previous profession]’, ‘assessment must be convenient’, ‘ensures that students understand concepts’, and ‘to make sure that the students understand what the subject is all about’. Declarative statements about self-perception were also very helpful in unravelling the participants’ implicit theory-in-use, as were statements related to their perceptions of assessment. The three quoted statements that follow sum up the Quantitative, Behaviourist assumptions (Biggs 1995) that most participants had about assessment: ‘grading students, measuring student knowledge [declarative], assessing student suitability for a chosen course, assessing that student understanding of concepts...
is in line with the lecturer's thinking'. The phrase ‘assessing understanding’ featured many times but it was rarely used in a Constructivist context, leading to Biggs’s (1996) assertion that ‘all teachers say they teach for understanding, but few do so in any sustainable way’. The majority of participants admitted the need for continuous professional development in assessment practices, as long as it was not delivered in ‘lecture style’ but using a practical approach.

**Observed Outcomes of the Assessment Design Process**

The analysed assignment documents that had been created by the participants were classified according to Bigg’s (1995) six modes and contexts of assessment. Two could be classed as Quantitative-Standards-Decontextualized. These were examples of 1970s-style criterion referenced testing, also referred to as mastery learning. This mode is best suited for testing basic and behavioural skills. Their rooting in the quantitative framework limits their application to lower cognitive skills and a superficial approach to learning. Another document was classified in the Qualitative-Standards-Decontextualized mode. This is a developmental mode of assessment that focuses on the growth of skills and concepts and is suitable for assessing legitimate declarative knowledge but does not focus on applications of knowledge. It has a positive backwash since it encourages learners to think higher. One analysed assignment, which had not been written by any of the participants but had been used by one of them, was categorized as Qualitative-Standards-Situated and required students to create a portfolio of learning experiences. It was a type of performance assessment that was situated in a work-based context and was meant to assist the students’ development of skills in applying what they learned in real-world situations, by planning, implementing plans, reflecting on their strategies, and recording their experiences. On a final note, participants whose assessments were based on quantitative foundations tended to enact limited mental models. Heinrich (2017) described limited mental models as types who understand the value of assessment linked to proximal assessment outcomes and behaviours but who do not attempt to connect their work on individual and programme level assessment goals unless specifically asked to do so. Their assessment behaviour is often guided by high expectations to align assessment goals, data and decisions on behalf of program accreditation. On the other hand, the participants who based their assessments on emerging paradigms (Shepard 2000) could be described as having a connected mental model. These often report specific mentoring in applied assessment practice and gain assessment experience and expertise from their current environment rather than coursework in their disciplines (Heinrich 2017).

**Conclusion**

A constructivist approach to grounded theory methodology was used to construct a preliminary conceptual model that attempted to explain the cognitive and behavioural processes at work during the creation of assessment briefs by MCAST science lecturers. Contextual factors, which were grouped under the major categories of knowledge (assessment literacy), skills, and beliefs were found to have a direct bearing on various stages of the assessment design process. The findings revealed that most lecturers’ approaches were determined by assessment views rooted in the twentieth-century dominant paradigm, which is still the dominant discourse in higher education assessment, despite the body of evidence pointing to better practices aligned to constructivist learning theories. Strong mentoring and a forward-looking learning culture can help lecturers, especially those having a connected mindset, to make the leap into better ways of designing constructively aligned assessment.
Limitations of the Study

This preliminary study was carried out on a small number of participants due to the short time-frame available. Therefore, the theoretical saturation required to build a comprehensive, explanatory theory was not achieved. The conceptual model that resulted from the analysis of the emerging theoretical constructs can be refined and improved upon by gathering more data through theoretical sampling to saturate and explain fully the major conceptual categories, including their properties and dimensions. Theoretical sampling will involve the pursuit of pointers that arose from the analysis that will enable the researcher to build a more comprehensive picture of patterns, gaps, and interrelations between concepts within the emerging conceptual framework (Birks and Mills 2015).

Implications for Stakeholders

Wiliam (2010) asserts that research in the UK has shown that, even though the millions of pounds were spent to buy more effective textbooks and to saturate classes with computers and high-tech interactive whiteboards, these measures have had a minimal impact on student achievement. Even class-size reduction and setting by ability have had only marginal improvement effects. ‘The only thing that matters,’ affirms Wiliam, ‘is the quality of the teachers.’ A culture of continuous professional development is the key for the improvement of any educational institution and this entails a focus on the ‘small number of things’ that are likely to improve outcomes for students. Assessment is certainly one of those key improvement areas. The findings of this study coincide with Shepard’s (2000) conclusions that ‘dominant theories of the past continue to operate as the default framework affecting and driving current practices and perspectives’. Assessment practices based on Constructivism, which has become the dominant espoused theory in education (Biggs 1996), have been found to result in large improvements in student achievement and higher-order skill development. However, as this study has also revealed, in higher education, the emphasis still remains on assessment of learning rather than assessment for learning (Quinn in Clarence et al. 2015). The assessment design process model resulting from this study could therefore be useful for educational organizations to identify those variables that lie within their control and that can have an impact on the assessment approaches that members of their teaching staff take. It is clear that the one-off training-course model that simply exposes educators to knowledge of constructivist assessment practices (especially if delivered using traditional lecturing style) does not work, since it is itself based on the Behaviourist paradigm and will not lead to a change in enacted mental models of the participants. Extra-individual factors affecting ways assessment is practised is certainly one of those key improvement areas. The setting up of institute-based assessment learning communities, which discuss assessment practice, set targets, try out new assessment methods, and share their reflections upon their effect in class. Thus lecturers will be exposed to ‘a variety of living examples of implementation’ that will form a beaten path to encourage them on their way to changing their approach to assessment (Black et al. 2004).
Recommendations for further research

Due to the qualitative nature of this study, the findings cannot be generalised for all educators but were intricately related to the particular context in which the inquiry was carried out. Similar future studies in other contexts, such as other higher education institutions and compulsory schooling institutions, could be carried out to build a more comprehensive picture of assessment practices nationwide. In addition, another study could be carried out at MCAST to investigate how the effects of assessment practices are perceived by the students in order to gauge their effects on their learning and motivation. The suitability of the assessment design process model to explain the assessment approaches taken by educators in other contexts could also be tested in future studies.

References


