Journal of University Teaching & Learning Practice

Volume 16 | Issue 3 Article 5

2019

An online testing design choice typology towards cheating threat minimisation

Albert Munoz
University of Wollongong, amunoz@uow.edu.au

Jonathon Mackay
University of Wollongong, jlm981@uowmail.edu.au

Recommended Citation

Munoz, A., & Mackay, J. (2019). An online testing design choice typology towards cheating threat minimisation. *Journal of University Teaching & Learning Practice, 16*(3). https://doi.org/10.53761/1.16.3.5

An online testing design choice typology towards cheating threat minimisation

Abstract

Online testing is a popular practice for tertiary educators, largely owing to efficiency in automation, scalability, and capability to add depth and breadth to subject offerings. As with all assessments, designs need to consider whether student cheating may be inadvertently made easier and more difficult to detect. Cheating can jeopardise the validity of inference drawn from the measurements produced by online tests, leading to inaccurate signals and misperceptions about what students know and can do. This paper extends theoretical understanding about cheating behaviours to link online test design choices and their influence on a student's capability and willingness to cheat. This research reviews the literature on cheating theories and a typology construction methodology to relate common online test design choices to their cheating threat consequence. In doing so, the typology offers educators designing online tests general normative guidance aimed at reducing threats to assessment inference validity, and academic integrity in general, brought about by student cheating. While we admit that cheating cannot be completely eliminated in online testing, the guidance provided by the typology can assist educators to structure online tests to minimise cheating.

Keywords

Cheating, Test Design

Journal of University Teaching & Learning Practice

Volume 16 | Issue 3 Article 5

2019

An online testing design choice typology towards cheating threat minimisation

Albert Munoz
University of Wollongong, amunoz@uow.edu.au

Jonathon Mackay
University of Wollongong, jlm981@uowmail.edu.au

An online testing design choice typology towards cheating threat minimisation

Abstract

Online testing is a popular practice for tertiary educators, largely owing to efficiency in automation, scalability, and capability to add depth and breadth to subject offerings. As with all assessments, designs need to consider whether student cheating may be inadvertently made easier and more difficult to detect. Cheating can jeopardise the validity of inference drawn from the measurements produced by online tests, leading to inaccurate signals and misperceptions about what students know and can do. This paper extends theoretical understanding about cheating behaviours to link online test design choices and their influence on a student's capability and willingness to cheat. This research reviews the literature on cheating theories and a typology construction methodology to relate common online test design choices to their cheating threat consequence. In doing so, the typology offers educators designing online tests general normative guidance aimed at reducing threats to assessment inference validity, and academic integrity in general, brought about by student cheating. While we admit that cheating cannot be completely eliminated in online testing, the guidance provided by the typology can assist educators to structure online tests to minimise cheating.

Keywords

Cheating, Test Design

Introduction

The increased popularity of online course offerings has also brought educators new challenges in assessment design. Online testing via learning-management systems (LMS) are seen as a promising way to increase the efficiency of assessment processes (Russell 2010; Maier et al. 2016) with a broader range of instruments available to measure learning (Prince et al. 2009) in a cost-effective manner (Young 2001). As with most technology, new and innovative ways of assessment inevitably bring new avenues of risk, and the potential to make cheating easier (Harmon & Lambrinos 2008; Hylton et al. 2016). With an increase in the popularity of online assessment methods, academic-integrity research and practice have made inroads in the development of guidance to keep up with growing concerns about student misconduct. One such approach, applied to online testing, is to create designs that reduce students' opportunity and temptation to cheat.

Assessment design for cheating prevention is hindered by a combination of little guidance and large implications if assessments are poorly designed. This hindrance is especially likely to emerge considering the vast array of assessment types and the choices involved in their design. Recent works have proposed strategies to reduce the threat of cheating in online assessments (e.g. Ladyshewsky 2015), and ways to measure cheating tendencies across assessment designs (e.g. Bretag et al. 2018). Recent increases in the sophistication of contract cheating have generated practical guidance about how assessment-design choices can be manipulated to reduce it (Bretag et al. 2018; Harper et al. 2018; Bretag & Harper 2019). Despite these advances, guidance is scant relative to the sheer multiplicity of designs available to educators wanting to structure online assessments to minimise cheating threats. There are several methods for delivering assessment online, such as e-portfolios, augmented reality, immersive environments and learning analytics (Guerrero-Roldán & Noguera 2018). In many cases, design of online assessments will be constrained by the assessment type itself, and thus specific guidance to inform the design process for one type of online assessment may be irrelevant for another. Given the broad range of online assessment types, this research will focus on the use of online tests comprising a number of assessment items that question and examine the student's competencies by eliciting responses to problems or questions. A further exploration into the literature on online test design uncovers the same issue of scant guidance relative to the large number of potential designs. Furthermore, the links have not yet been determined between how individual online test design choices and their cheating-threat consequence – individually and holistically – affect cheating.

There are also several methods to measure cheating threats induced by online test design choices (e.g. human experimentation, self-reporting surveys, data mining and profiling of online behaviour). However, many of these methods are affected by issues relating to ethical treatment of humans, as in the case of testing assessments by segregating students into control and experimental assessment groups (McNeill 1993; Hansen & Stephens 2000), validity threats of self-reporting misconduct (Ladyshewsky 2015) or potential student mis-profiling from a lack of relevant contextual data (Sander 2016). Given the existence of these issues, the current research employs theories explaining cheating behaviour to inform online test choice threat implications.

The assessment-design process produces a metric (i.e. a test) that can measure what students know and can do (Taras 2005; Trumbull & Lash 2013; Baird et al. 2017). In formative assessments, these measurements can be used to provide feedback to students as a means of enhancing learning (Ramaprasad 1983; Wiliam & Thompson 2017), while in summative assessments, the measurements can be used for the purposes of certification (Boud & Falchikov 2007). Thus, poorly designed online tests can produce misleading measurements about student performance,

misinforming educators, and ultimately undermining academic integrity. If assessments are to serve the goals of education, then assessment design must consider student cheating as a real validity threat to assessments as instruments of learning measurement.

This research uses a typology-construction methodology to link online test design choices and threat implications to academic integrity. Generally, typologies seek to classify complex phenomena into relevant dimensions and establish relationships (Doty & Glick 1994). Typologies are often conceptual in nature (Bailey 1994); however, they often serve as important tools towards theoretical development (Doty & Glick 1994; Collier et al. 2012). Accordingly, this research uses extant literature and theories to develop a typology that can be used as a model towards guiding assessment design. Because it is a multidimensional model, the value of the typology lies in its conceptualisation; that is, the linking of cheating theories to design choices. The alternative of proposing a large set of hypothetical designs is attractive, and an important area of future research that necessitates a strong theoretical foundation to be rigorous.

To date, no studies report comprehensive connections between online test design choices and cheating threats. This research addresses scholars' calls for educators to adopt a holistic approach to addressing issues of academic integrity in educational institutions (Bretag 2013), by carefully structuring online tests to reduce the utility of students engaging in cheating (Lancaster et al. 2016). Although no assessment can be designed to perfectly deter all cheating (Bretag & Harper 2017), this research argues that a more comprehensive understanding of design choices and their relationship to cheating behaviours can enhance online test design practice. To this end, this study makes inroads to inform the online test design process towards greater robustness to cheating and reduce validity threats to the inference drawn from such measurements.

Assessment as measurement

Scholars argue that what a student learns from instruction is impossible to predict, and that assessments may be the key to effectively measuring student understanding (Wiliam 2011). Assessments produce measurements about what people know and can do (Baird et al. 2017) in a learning process (Keeves 1994; Reeves & Hedberg 2007; Gikandi et al. 2011). Assessments in education, as any measure of performance, is ultimately founded in measurement theory (Trumbull & Lash 2013), which is primarily concerned with the assignment of values to phenomena of interest (Stevens 1959; Suppes & Zinnes 1962). Measurements themselves constitute models (or simplifications) of complex phenomena for the purposes of gaining a greater understanding of the phenomenon of interest (Hand 2004). In the context of online testing, test grades are measurements that constitute evidence used in the creation of inference; these evidences help an educator infer what a student knows and can do (Pellegrino et al. 2001; Trumbull & Lash 2013).

The design of online tests as metrics requires some means of assessing the metric itself. Measurement theory points to three primary criteria: reliability, validity and uniqueness (Hand 2004). These three conditions are necessary to enable robust evaluation of choices and provide valuable feedback during the design process. *Reliability* relates to the scale of error captured by the metric across measurements (Hand 2004). To ensure reliability the measures must be invariant to changes that do not represent aspects of the phenomenon of interest and variant to changes in the aspects of the phenomenon that the metric is intending to describe (Hand 2004; Suppes 2009). Specific to online tests, reliability is the degree to which what is assessment is dependable in measuring the level of knowledge being developed (Gikandi et al. 2011). In the presence of

cheating, as additional error is captured by the online test, the reliability of an assessment will decrease.

Validity is the extent to which a metric measures what it is intending to measure (Hand 2004). In assessment practice, validity is related to the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of how the test scores facilitate the generation of inference (Messick 1995). Research into the concept of assessment validity highlights the importance of investigating the consequences of assessment use, and subsequent impacts on the educational system (Moss 1992). Views of assessment validity point to the inferences made from the importance of the concept of online test results in the evaluation of test designs (Gikandi et al. 2011). The threat facing many educators using online tests is that students may present misleading evidence, inciting erroneous inference production by the educator, especially if cheating is undetected. A more comprehensive view of validity in educational assessment envelops the concept of score interpretation and use (Messick 1995). Validity threats also have social consequences, as tests with inherent weaknesses (e.g. disproportionate allowances for students to cheat) have consequences for the greater contribution of students to society. Both reliability and validity are inextricably linked in situations where cheating actions threaten the validity of the assessment (Knight 2002).

The concept of metric *uniqueness* relates to the degree to which the chosen representation approaches being the only one possible (Hand 2004), while still remaining logically consistent (Pedhazur & Schmelkin 2013). In any measurement exercise, it is important to discriminate between alternative measures of a given phenomenon (Suppes 2009) and recognise that metrics will have unequal utility (Luce & Suppes 2002; Hand 2004). When tests are adapted for use in a particular assessment space, variations in online test design are known to lead to differences in which learning is measured (Tian 2007), as well as uncovering differences in performance among cohorts (Bridges et al. 2002). If an online test design that is robust to cheating is desired, uniqueness will be characterised by a design's ability to reduce cheating occurrences. Thus, the uniqueness of such assessments requires sufficient knowledge about cheating and its relationship to online test design choices to guide the design processes and preserve the effectiveness of measurements.

Cheating

Cheating refers to a vast array of actions students can undertake to gain an unfair advantage, which as a consequence may alter perceptions of student ability (King et al. 2009); these can include unauthorised use of resources within stipulated boundaries (Khare & Lam 2008). Of course, the range of actions students can undertake far exceeds the influence of assessment design. For example, students may choose to intrusively undermine the assessment itself (e.g. hacking) (Dawson 2016; Ullah et al. 2016). These actions can target weaknesses in the LMS, and are typically the responsibility of the LMS administrator, rather than the educator tasked with assessment design. To this end, this research limits the scope of cheating to those actions that are non-intrusive and are logically influenced by the online test design.

Logically, if online test design is to be aimed at reducing cheating behaviours, important antecedents to cheating behaviours must be included in design considerations (Bolin 2004). Reviews of the literature uncover two theories explaining cheating behaviour: the general theory of crime and cognitive offloading. The general theory of crime (GTC) cites a lack of self-control and desire to pursue immediate desires as an explanation for why people commit crimes (Gottfredson & Hirschi 1990). Applications of GTC to cheating in assessments by Bolin (2004)

explain cheating as the product of attitudes towards academic dishonesty – further linked by Mouberry (2009) to motivational theory – and the perceived opportunity to cheat. The attitudes towards academic dishonesty and the motivational forces can be characterised as either intrinsic goals (i.e. linked to personal mastery and increasing one's competence) or extrinsic goals (i.e. performance and grades) (Whitley 1998; Jordan 2001; Rettinger et al. 2004). In conjunction with these motivations, the student has to perceive an opportunity to cheat to facilitate or enable cheating behaviour (Davis et al. 1992).

A second theoretical foundation stems from the recognised limitations of the human mind (Simon 1982), and the tendency to delegate cognition to technological or other aids (such as phones for remembering numbers and appointments) to overcome capacity limitations – referred to as cognitive offloading (CO) (Risko & Gilbert 2016). When individuals are tasked with performing something that has overwhelming cognitive demands, they tend to take action to alter the requirements of a task and reduce cognitive demand (Risko & Gilbert 2016). Humans choose to adopt offloading strategies that are perceived as subjectively optimal – even if they could be judged by others to be suboptimal – when internal costs of effort are taken into account (Anderson 1990). For example, humans use digital technology instead of committing information to memory as a form of cognitive offloading; if the information is going to be available via the internet, there is a lower perceived need to commit the information to memory (Storm et al. 2017). Thus, effort perceived to be beyond cognitive capacity is avoided in pursuit of minimisation of cognitive workload (Reichle et al. 2000). In the context of assessments, the cognitive demands that assessment items place on a student can result in offloading action if the benefits of offloading are considered superior, especially when the demand being offloaded is considered unnecessary with respect to learning goals (Risko & Gilbert 2016).

Together, these theories allow for a broad view of the human characteristics that lead to cheating behaviours, with wilful intention coupled with opportunity as one cause, and a tendency to engage external resources as another. GTC and CO explain cheating from a perspective that informs online test design. From the perspective of GTC, online tests should aim to manage both the motivations to cheat (i.e. reducing pressure to achieve intrinsic and extrinsic goals) and the available opportunities to cheat. Meanwhile, CO-driven online test design should endeavour to manage the cognitive demands that the assessment items place on a student, lest they be perceived or anticipated as being greater than the benefit of offloading. In these cases, educators seeking to minimise cheating threats in online tests would intuitively seek to account for a design's influence on such student behaviours. However, practitioners could argue that this theoretical guidance alone is insufficient to guide online test design. Therefore, we use these theories to justify the relevant dimensions of online test design choices aimed at cheating minimisation. Specifically, the theories help populate the typology with links between online test design choices and guidance to use these choices to reduce motivations, opportunities to cheat, as well as manage test items' demand on student cognition.

Cheating represents a corruption to learning measurements, as it can threaten the validity of inferences drawn from online tests; thus the nature of the test can be linked to a threat magnitude. This relationship is supported by risk-management scholars, who stress that an evaluation of the choices made in terms of their impact upon the desired behaviour are important (Haimes 1991; Haimes 2015), and need to be evaluated using logically sound, practical criteria (Fischhoff et al. 1980). For educators tasked with making choices about how online tests are designed, this implies that both individual choices and combinations of these choices have an influence on the probability of student cheating.

Cheating and online test design

An online test design can simply be an equivalent translation of a paper-and-pencil test and yield similar results, especially when the testing environment is similar (e.g. proctored settings) (Bugbee Jr 1996). However, online tests offer far more flexibility than can be exploited in the crafting of new designs. These new designs must include some elements that aim to mitigate the risks of cheating (Ladyshewsky 2015; Bennett et al. 2016). With these risks in mind, perceptions of increased cheating in online environments have resulted in the realisation that transitions to online assessments require some redesign (Hollister & Berenson 2009). Redesign of assessments for an online environment should be directed at preserving the validity of the assessment of learning-measurement instruments. Such redesign can be daunting to educators (Zanjani et al. 2017), and they should be offered considerable guidance in how to minimise cheating threats (Bolin 2004).

Within the realm of online test design, there are a multitude of individual suggestions to prevent cheating (e.g. issuing warnings, setting time limits, monitoring), each with their own expected influence on the threat of cheating. Specific to online test design, extant guidelines suggest a reasonable and fixed amount of time to respond to each item to preclude students from having time to source answers from surrogates (Schultz et al. 2008). Furthering this strategy, Ladyshewsky (2015) recommends permitting only one item to be viewed at a time (randomly drawn from a large pool of items), setting time limits per question and preventing backtracking to prevent students who finish early from going back and copying items with the intention of sharing. Moreover, the literature encourages items that demand critical thinking rather than knowledge regurgitation (Harper 2006). The amount of time students have access to the online test is also a choice, and limiting the time the assessment is available also limits the time students have to engage in cheating (Ladyshewsky 2015). Such guidance, although valuable in its own right, assumes that most online test design choices are made independently and driven by practitioner experience. Moreover, these recommendations are often made in isolation of each other, with few articles accounting for any combinatorial effects. One example is the work of Sewell et al. (2010), who argue that in situations where exams have high stakes (>20%), students should be proctored. As with any identified risk, multiple actions inevitably lead to trade-offs, which can result in complementary or substituted effects on desired outcomes (Haimes 2015). The combined effects of multiple online test design choices can also dramatically increase complexity and jeopardise the clarity of normative guidance, or may inadvertently lead to designs that encourage cheating.

The environment in which testing takes place can also be influenced. For example, the adoption of proctoring is a popular approach to monitoring the environment in which students are assessed. Scholars believe that cheating in online tests occurs more frequently in unproctored environments, and associate proctoring with discouraging student cheating (Hollister & Berenson 2009; Prince et al. 2009). In online assessments, web-based cameras (webcams) are proposed as a technology-based solution to enable proctoring and ultimately deter cheating (Hylton et al. 2016). Other applications of online testing have conducted tests in a proctored environment, prohibiting the use of removable hard drives and disabling access to the internet outside the LMS (Ghauth & Abdullah 2010). Proctoring allows both identity validation and the monitoring of assessment behaviour, and is supported by the idea that students performing assessments in unsupervised environments may cheat more than those in proctored settings.

The intuition of many educators (for example, Harmon & Lambrinos 2008) aligns well with assertions in the literature that online tests are more vulnerable to cheating than offline counterparts. In contrast, some research suggests that such assertions do not appear to have conclusive evidence that online tests have led to an increase in cheating (Underwood 2006;

Ladyshewsky 2015). One possible source of this disagreement are criticisms about methodological choices. Some studies have relied on self-report surveys to assess the extent to which cheating occurs, which may not yield valid results even with assurances of anonymity (Finn & Frone 2004; Styron & Styron Jr 2010; Fask et al. 2014). Irrespective of which side of the argument is correct, no testing method can conclusively eliminate all cheating. To avoid potential conflict over methodological choices, this paper relies on a review of the underlying theories that explain cheating. These theories, and their components, serve as a context from which a typology can be developed. This paper posits that the existence of theories explaining cheating provide a solid foundation for developing comprehensive normative guidance to support online test designs that minimise cheating threats.

A typology for online assessment design

Typologies classify and organise phenomena according to a specific set of dimensions or variables (Bailey 1994; Collier et al. 2012). They offer practical implications through establishing criteria that assist in identifying patterns and relationships, which in turn serve as tools towards theory-building and practice (Doty & Glick 1994; Collier et al. 2012). For example, Higbee and Thomas (2002) proposed categorisations of the perceptions of what constitutes cheating, based on first-level experiences expressed as constructed types; they uncovered considerable variance in perceptions of cheating. Typologies also allow researchers to develop and refine concepts into conceptual models that can be deployed for empirical research and practice (Collier et al. 2012). The set of types forms a model of the phenomenon of interest, which may be too complex to describe in the entirety of its detail. Thus, the aim of the typology is ultimately to enhance understanding and further the creation of knowledge by including all the important characteristics for the purposes of the research, while excluding all those that are less important. In the context of this research, the typology of online test design choices is constructed specifically to include relationships between choices and their respective cheating threat, while excluding online test choices to which cheating threat is insensitive.

Conceptually, an assessment places a demand upon the student's cognition to produce an answer. From a CO perspective, a student then can choose to produce the answer or obtain it from another source (Risko & Gilbert 2016). Furthermore, if the student perceives the demand to be greater than their cognitive capacity, they may alter the information-processing requirements of a task, otherwise known as engaging in cognitive offloading (Risko & Gilbert 2016). The student's perceived benefit from offloading will also be contingent on the fact that the demand being offloaded is unnecessary with respect to the learning goal (Paas et al. 2003; Risko & Gilbert 2016). From a GTC perspective, a student's attitude towards academic integrity, motivations and desires to perform well in assessments may not be well aligned, and any opportunity to cheat will influence the choice to cheat (Gottfredson & Hirschi 1990; Bolin 2004; Mouberry 2009). These two theories can be reconceptualised as two forces, each dominating the ends of a continuum and together influencing the choice to cheat in online tests. At one end are human tendencies associated with identifying and engaging surrogates to complete assessment tasks. At the other end are assessment-design choices associated with the opportunity to cheat, moderated by the strength of motivations. The typology is organised by first providing a description for each online test design choice, starting with those that influence cognitive offloading. As the dominant theory shifts towards GTC, motivations and attitudes towards academic integrity become a moderating force, and ultimately drive the online test design choice to reduce the opportunity to cheat. Each dimension is linked to a choice and to cheating-threat implications, and described at its extremes (Table 1).

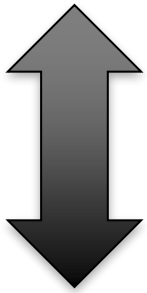
Forces	Dimension	Assessment-design choices	Implications to cheating threats
Cognitive offloading	Anticipated cognitive demand against the reward of achievement	Level of cognitive demand placed on student	Greater demand on cognition leads to offloading.
	Expected effectiveness of offloading	Producing test items with answers not readily obtainable via unauthorised sources (e.g. surrogates)	Capacity to obtain acceptable test item responses from unauthorised sources increases the effectiveness of offloading.
	Stakes per assessment item	Quantity of online tests and stakes per test	Higher stakes increase pressure placed on students, leading to greater propensity to cheat.
Motivation to cheat	Number of assessment items in allotted time	Quantity of items in the test, time allotted to conduct test	Too many items in too short a period could exceed cognitive load threshold and increase the chance of offloading, as well as increase feelings of unfairness and motivate cheating. Too few items in too long a period could increase the effectiveness of offloading.
	Perishability of assessment items	Asynchronicity or synchronicity of test; use of static item sets or randomised items drawn from item pool	Asynchronous testing risks the passing on of items and responses to students taking assessments later. Static item sets used repeatedly risk students identifying patterns and disseminating items and responses.
	Proctoring	Use of proctoring technology	Invigilation of students during tests decreases the risk of cheating.
Capability to cheat			

Table 1. Online test design choice typology and implications for cheating threat

Anticipated cognitive demand against the reward of achievement

Educators tread a fine line when designing online tests, as a balance must be struck between setting tests that fail to measure learning by setting too low a demand on cognition and those that overload cognition. Of course, extrinsic motivation can enhance the utility of completing assessments without offloading and can serve to push the boundary of cognitive demand to tests that are cognitively demanding while still presenting sufficient reward of achievement to students that counters the motivation to offload. Thus online test design choices must determine the level of cognitive demand placed on students, acknowledging and anticipating the effort avoidance that underpins CO. Effort avoidance is influenced by a subjective evaluation of efforts (Clithero & Rangel 2013), and is conceptualised as an anticipated effort associated with action in the decision-making process (Dunn et al. 2017). Online test design needs to account for both the cognitive demand, balanced against the reward of achievement (i.e. intrinsic motivation), and the anticipated effort expectations constructed by students during the education process.

Expected effectiveness of offloading

Alternative, unauthorised sources can be employed in several forms, each with utility and expected benefit. This paper expresses this benefit in terms of the effectiveness of offloading as influenced by online test design choices. For example, students can use search engines to find answers to items in an online test (Ravasco 2012). This process involves students evaluating how well search-engine results will serve as responses to the item, with higher utility being given to answers that respond well to the item. In the context of online tests, access to the broader internet can be easily gained and exploited. Online tests that use items extracted from textbook publishers' question banks present some difficulties, as they are often available online. Students doing online searches while undertaking unproctored online tests can often readily find correct responses. The access to additional information influences how students think, governs their decisions about their knowledge, and influences whether they are willing to volunteer answers or seek other means of answering questions (Ferguson et al. 2015). One potential solution is to adopt customised item sets, as online searches for such items will reduce the effectiveness of offloading.

Multi-party cheating involves a collective effort, most notably through collusion (Perry 2010) and contract cheating (Lancaster & Clarke 2007; Rigby et al. 2015). Students considering engaging in multi-party cheating evaluate how useful it would be to engage surrogate parties capable of performing better than the individual (i.e. extrinsic motivation). The availability of such surrogates can be restricted by limiting access times, for example. However, student evaluations about offloading effectiveness are likely to be complex, with other dimensions of the typology having some significance. Generally, online test designers should bear in mind that students contemplating cheating are likely to perform some evaluations about the effectiveness of offloading, and limiting the availability of surrogates is an important consideration.

Stakes per assessment

"Assessment stakes" refers to the value a particular test has towards a student's final grade, typically represented as a percentage value of the overall total (Au 2007). The scale of stakes ranges from "high stakes" tests, where those assessed aim to conceal ignorance or suggest competence, and "low stakes" assessments, where learners react more openly about their limitations (Knight 2001). The relationship between stakes and cheating threats is well known in the literature. High-stakes assessments drive changes in behaviours (Au 2007), tend to encourage higher frequency of cheating (Park 2003) and should be avoided (Carstairs & Myors 2009). In contrast, low-stakes assessments carry with them less pressure (Holmes 2015), leading fewer

students to engage in cheating behaviour. Low-stakes online tests are popular, partly because they are thought to promote consistent work over time (Bennett et al. 2016). Generally, the literature tends to unilaterally recommend a lowering of stakes to reduce cheating threats.

Conducting tests online facilitates greater scale and allows students to engage with a larger number of lower-stakes tests, but important practical limitations do exist. As the number of tests increases, the stakes per test decrease. However, a multitude of other models exist, with institutional policy and practical constraints playing a role in these distributions. For example, some models of stakes' distribution will design multiple instances of formative assessments with relatively low stakes, with a summative assessment at the end of the term having considerably higher stakes. If delivered as online tests, each high- and low-stakes design should differ to account for their respective cheating threat.

Number of assessment items in allotted time

Timed tests decrease the opportunity to cheat (Rovai 2000). Conversely, assessments where time is unlimited allow for greater engagement of surrogates. This general rule, unlike stakes per assessment item, may be called into question if the time required to respond to an item far exceeds the time allotted. As a hypothetical example, an online test could require students to complete 100 questions, each taking about 30 seconds to complete, but the students could only be given five minutes to complete the test. The insufficient time allotted would be likely to drive students to cheat, as attempting to complete the assessment might be perceived as futile unless they engaged in offloading. Such feelings of futility feed into motivational forces to cheat, as student perceptions of their ability diminish (Murdock & Anderman 2006).

Conversely, providing too much time to students to complete an online assessment might provide too much flexibility, and students might engage in cheating to ensure that their performance is maximised. Furthermore, motivations to cheat are known to be increased if the incurred cost of engaging in cheating is reduced (Murdock & Anderman 2006). This thought experiment uncovers a nonlinear relationship between the number of items in a test and the time provided to complete the test present at the intersection between motivations to cheat and cognitive offloading. Thus, design choices about online quizzes where too much or too little time relative to the number of assessment items is provided can lead to an increase cheating threat.

Perishability of assessment items

Assessment-design models may reuse test items across online test instances. However, this premise is not universally applicable, as items may not all have the same prospects for reuse (Oliver 2006). As items are reused, they may become partially or entirely unfit for use. Items within an online test can be viewed as an inventory that can undergo change in its effectiveness to demand knowledge from students in a way that preserves academic integrity. Such items can be considered perishable (consistent with the definition of perishable inventories by Nahmias 1982). Although the lifetime of items is typically unknown, the perishability characteristic of an item has design consequences. Ultimately, there is no complete solution to the perishability problem, but some choices may help. Two choice options exist for online test design: randomisation of items (Fluck et al. 2009; Arnold 2016) and synchronicity of online test delivery (Chao et al. 2012).

Randomisation of item sets has a considerable influence on students' ability to identify (and potentially exploit) pattern recognition. Humans have an inherent ability to discern patterns from data, with repetition being one important enabler to pattern identification (Eysenck & Keane 2015). In the context of online tests, the use of a static question set will presumably make cheating

easier, as students who recognise a pattern can share the code or respond to items without relying on the knowledge acquired. One way to avoid students' recognising patterns is to use a large pool of items to produce unique tests for each student (Rovai 2000). However, if the pool of items is too large, online tests may sample items in a manner that may not be broad enough to adequately assess all the desired content. To counter, designs may partition item banks to specific areas of content, and sample items from each partitioned item set as necessary. Another issue in the production of large item banks, especially considering the aforementioned guidance of employing customised item sets, is the existence of practical time constraints.

Additionally, LMSes provide options to conduct online assessments synchronously (for all students at the same time) or asynchronously (Chao et al. 2012). Students taking online asynchronous assessments are required to finish the assessment procedures either in their own time or within a specific duration, but without any real-time interaction between the instructor and learner during the assessment activity (Chao et al. 2012). Relative to asynchronous online test designs, synchronous designs will minimise cheating in online tests, as the entire cohort is presented with a set of assessment items that cannot be passed on to other students; the pool of potential multi-party cheating surrogates is limited, as most potential surrogates will be engaged in the same test; and the relatively fewer opportunities to identify patterns among assessment items will lessen item perishability. An alternative asynchronous solution to minimise cheating could rely on further expanding item banks to generate independent sets of randomly sampled items from different item pools. This would reduce the likelihood of patterns identification, and students taking the same test at a later time would be less likely to cheat. Of course, the optimal choice would be to design online tests that use random sampling of items synchronously (Marks & Cronje 2008; Fluck et al. 2009). However, if such a design is infeasible, online test design needs to consider that assessment-item perishability is exacerbated by ease of pattern recognition and surrogate availability.

Proctoring

Proctored testing involves the use of invigilators to oversee students whilst completing a test that, amongst other support roles, serves to detect cheating. The use of proctoring also enhances identity security and can increase the detection of other instances of academic dishonesty (Rovai 2000). The adoption of high degrees of proctoring comes with some time and cost constraints (Trenholm 2007). The proctoring of online tests is difficult to implement (Fask et al. 2014), although webcam-based proctoring is reported to significantly deter users from misconduct (Hylton et al. 2016). Another form of proctoring can also be done remotely, by removing the student's computer's capability to search online (DeMara et al. 2016). Therefore, educators seeking to include proctoring in online tests face issues with the ability to implement proctoring against institutional resource constraints.

Summary and discussion

With the growth of online learning environments, educators are finding more-accurate ways to assess student learning (Styron & Styron Jr 2010). However, online test design choices need to be accounted for, lest they motivate the wrong behaviours (Baird et al. 2017). Empirical research has proven that the relationship between learning and the ways learning outcomes are assessed matters (Tian 2007). Alternative test methods provide different measures about what a student has learned (Scouller 1998, Bridges et al. 2002), consistent with the view that not all measurement methods are equally variant to a phenomenon of interest (Suppes & Zinnes 1962). We applied this argument to the choices made in the design of online tests and their threat consequence, specifically to student cheating.

To inform online test design, the research grounded student cheating via two theories: the general theory of crime (Gottfredson & Hirschi 1990) and cognitive offloading (Risko & Gilbert 2016). The theoretical understanding of cheating allowed a typology to be constructed that relates theories of student cheating to assessment-design choices. The typology provides normative guidance for online test design and assessment of cheating threat, with the intended aim being to reduce motivations, opportunity and propensity to engage in cognitive offloading. The typology, by itself, is not a prescriptive roadmap to optimal online test design; rather it informs practitioners of the potential cheating consequences of particular design choices. The typology also lays the foundation for online test design evaluation. LMSes commonly have the capability to act as passive data-collection repositories. Such data, analysed appropriately, can serve as an important feedback mechanism for improving test designs in their ability to inhibit cheating, uncover cheating patterns and provide real-time feedback to designers, allowing for rapid test-design enhancements.

Among the many challenges of assessment design, a persisting concern is that of conducting the assessments while maintaining academic integrity (Hollister & Berenson 2009). Cheating threatens the validity of the assessment task as an instrument to measure student learning (Pellegrino et al. 2001), and cheating in online tests can involve a myriad of strategies (Ladyshewsky 2015). The guidance presented in this study about individual online test design choices and cheating threat provide an important first step for many. This study focused on guidance at the level of individual choice, acknowledging that the large number of choices – and options for each choice – can result in a very large number of designs. We thus purposely avoid providing sample designs or templates for online tests, as the sheer number of possibilities would be too large to describe in one paper. However, we are confident that the guidance presented, when applied to the online test design process, will enable educators and assessment designers to minimise cheating threats. Taking the assumption that the volume of strategies to cheat is likely to increase, a dynamic perspective of cheating threats leaves much work to do for those aiming to preserve the validity of online tests into the future. Furthermore, advancements in LMSes and online assessment platforms should yield a greater variety of online assessments into the future. The guidance provided in this paper is applicable to a broader range of online assessments. Further research into the nature of other online assessments and theoretical "stress-testing" could be applied to further enhance the typology.

A limitation of this research that has been noted in the literature on online test design is a lack of accounting for interactions among design choices, particularly as it relates to cheating threats. The combinatorial effects of multi-choice designs can lead to instances in which one choice increases, decreases or negates the effect of another choice on student cheating. Our research presents some of the more straightforward impacts on cheating, but is limited in the scope and scale of analysis. A comprehensive multi-choice extension to the typology that includes inter-choice interactions would further enhance the guiding value of the research towards the practice and evaluation of proposed designs. A deeper understanding of these combinatorial effects could lay the necessary groundwork for the development of robust design portfolios that could safeguard online tests against the threats of student cheating.

References

Anderson, JR 1990, *The adaptive character of thought*, Psychology Press, New York.

Arnold, IJ 2016, 'Cheating at online formative tests: Does it pay off?', *The Internet and Higher Education*, vol. 29, pp. 98-106.

Au, W 2007, 'High-stakes testing and curricular control: A qualitative metasynthesis', *Educational researcher*, vol. 36, no. 5, pp. 258-267.

Bailey, KD 1994, *Typologies and taxonomies: an introduction to classification techniques*, Sage Publications, Thousand Oaks, CA.

Baird, J-A, Andrich, D, Hopfenbeck, TN & Stobart, G 2017, 'Assessment and learning: Fields apart?', *Assessment in Education: Principles, Policy & Practice*, vol. 24, no. 3, pp. 317-350.

Bennett, S, Dawson,P, Bearman, M, Molloy, E & Boud, D 2016, 'How technology shapes assessment design: Findings from a study of university teachers', *British Journal of Educational Technology*, vol. 48, no. 2, pp. 672-682.

Bolin, AU 2004, 'Self-control, perceived opportunity, and attitudes as predictors of academic dishonesty', *Journal of Psychology*, vol. 138, no. 2, pp. 101-114.

Boud, D & Falchikov, N 2007, *Rethinking assessment in higher education: Learning for the longer term*, Routledge, Oxford.

Bretag, T 2013, 'Challenges in addressing plagiarism in education', *PLoS medicine*, vol. 10, no. 12, p. e1001574.

Bretag, T & Harper, R 2017, 'Assessment design won't stop cheating, but our relationships with students might' *The Conversation, Australia*, viewed 29 April 2019, viewed at <https://theconversation.com/assessment-design-wont-stop-cheating-but-our-relationships-with-students-might-76394>.

Bretag, T & Harper, R 2019, 'Contract cheating and assessment design', viewed 29 March 2019, <https://cheatingandassessment.edu.au/>.

Bretag, T, Harper, R, Burton, M, Ellis, C, Newton, P, Rozenberg, P, Saddiqui, S & van Haeringen, K 2018, 'Contract cheating: A survey of Australian university students', *Studies in Higher Education*, vol. no. pp. 1-20.

Bretag, T, Harper, R, Burton, M, Ellis, C, Newton, P, van Haeringen, K, Saddiqui, S & Rozenberg, P 2018, 'Contract cheating and assessment design: exploring the relationship', *Assessment & Evaluation in Higher Education*, vol. 44, no. 5, pp. 1-16.

Bridges, P, Cooper, A, Evanson, P, Haines, C, Jenkins, D, Scurry, D, Woolf, H & Yorke, M 2002, 'Coursework marks high, examination marks low: discuss', *Assessment & Evaluation in Higher Education*, vol. 27, no. 1, pp. 35-48.

Bugbee Jr, AC 1996, 'The equivalence of paper-and-pencil and computer-based testing', *Journal of research on computing in education*, vol. 28, no. 3, pp. 282-299.

Carstairs, J & Myors, B 2009, 'Internet testing: A natural experiment reveals test score inflation on a high-stakes, unproctored cognitive test', *Computers in Human Behavior*, vol. 25, no. 3, pp. 738-742.

Chao, KJ, Hung, IC & Chen, NS 2012, 'On the design of online synchronous assessments in a synchronous cyber classroom', *Journal of Computer Assisted Learning*, vol. 28, no. 4, pp. 379-395.

Clithero, JA & Rangel, A 2013, 'Informatic parcellation of the network involved in the computation of subjective value', *Social cognitive and affective neuroscience*, vol. 9, no. 9, pp. 1289-1302.

Collier, D, LaPorte, J & Seawright, J 2012, 'Putting typologies to work: Concept formation, measurement, and analytic rigor', *Political Research Quarterly*, vol. 65, no. 1, pp. 217-232.

Davis, SF, Grover, CA, Becker, AH & McGregor, LN 1992, 'Academic dishonesty: Prevalence, determinants, techniques, and punishments', *Teaching of Psychology*, vol. 19, no. 1, pp. 16-20.

Dawson, P 2016, 'Five ways to hack and cheat with bring-your-own-device electronic examinations', *British Journal of Educational Technology*, vol. 47, no. 4, pp. 592-600.

DeMara, RF, Khoshavi, N, Pyle, S, Edison, J, Hartshorne, R, Chen, B & Georgiopoulos, M 2016, 'Redesigning computer engineering gateway courses using a novel remediation hierarchy', Proceedings of American Society of Engineering Education 123rd Annual Conference and Exposition, New Orleans, LA, 26-29 June, viewed 29 April 2019, <https://peer.asee.org/26063.pdf>.

Doty, DH & Glick, WH 1994, 'Typologies as a Unique Form of Theory Building: Toward Improved Understanding and Modeling', *Academy of Management Review*, vol. 19, no. 2, pp. 230-251.

Dunn, TL, Koehler, DJ & Risko, EF 2017, 'Evaluating Effort: Influences of Evaluation Mode on Judgments of Task-specific Efforts', *Journal of Behavioral Decision Making*, vol. 30, no. 4, pp. 869-888.

Eysenck, MW & Keane, MT 2015, *Cognitive psychology: A student's handbook,* Taylor & Francis, New York.

Fask, A, Englander, F & Wang, Z 2014, 'Do online exams facilitate cheating? An experiment designed to separate possible cheating from the effect of the online test taking environment', *Journal of Academic Ethics*, vol. 12, no. 2, pp. 101-112.

Ferguson, AM, McLean, D & Risko, EF 2015, 'Answers at your fingertips: Access to the Internet influences willingness to answer questions', *Consciousness and Cognition*, vol. 37, no. pp. 91-102.

Finn, KV & Frone, MR 2004, 'Academic performance and cheating: Moderating role of school identification and self-efficacy', *Journal of Educational Research*, vol. 97, no. 3, pp. 115-121.

Fischhoff, B, Lichtenstein, S, Slovic, P, Keeney, R & Derby, S 1980, *'Approaches to acceptable risk: A critical guide'*, Decision Research, Eugene, OR, viewed 29 April 2019, <https://www.osti.gov/servlets/purl/5045395>.

Fluck, A, Pullen, D & Harper, C 2009, 'Case study of a computer based examination system', *Australasian Journal of Educational Technology*, vol. 25, no. 4, pp. 509-523.

Ghauth, KI & Abdullah, NA 2010, 'Measuring learner's performance in e-learning recommender systems', *Australasian Journal of Educational Technology*, vol. 26, no. 6, pp. 764-774.

Gikandi, JW, Morrow, D & Davis, NE 2011, 'Online formative assessment in higher education: A review of the literature', *Computers & education*, vol. 57, no. 4, pp. 2333-2351.

Gottfredson, MR & Hirschi, T 1990, *A general theory of crime,* Stanford University Press, California.

Guerrero-Roldán, A-E & Noguera, I 2018, 'A model for aligning assessment with competences and learning activities in online courses', *The Internet and Higher Education*, vol. 38, no. pp. 36-46.

Haimes, YY 1991, 'Total risk management', *Risk analysis*, vol. 11, no. 2, pp. 169-171.

Haimes, YY 2015, *Risk modeling, assessment, and management,* John Wiley & Sons, Hoboken, NJ.

Hand, DJ 2004, *Measurement theory and practice: The world through quantification,* Oxford University Press, Oxford.

Hansen, EJ & Stephens, JA 2000, 'The Ethics of Learner-Centered Education: Dynamics That Impede the Process', *Change: The Magazine of Higher Learning*, vol. 32, no. 5, pp. 40-47.

Harmon, OR & Lambrinos, J 2008, 'Are online exams an invitation to cheat?', *Journal of Economic Education*, vol. 39, no. 2, pp. 116-125.

Harper, MG 2006, 'High tech cheating', *Nurse Education in Practice*, vol. 6, no. 6, pp. 364-371.

Harper, R, Bretag, T, Ellis, C, Newton, P, Rozenberg, P, Saddiqui, S & van Haeringen, K 2018, 'Contract cheating: a survey of Australian university staff', *Studies in Higher Education*, vol. no. pp. 1-17.

Higbee, JL & Thomas, PV 2002, 'Student and faculty perceptions of behaviors that constitute cheating', *NASPA Journal*, vol. 40, no. 1, pp. 39-52.

Hollister, KK & Berenson, ML 2009, 'Proctored versus unproctored online exams: Studying the impact of exam environment on student performance', *Decision Sciences Journal of Innovative Education*, vol. 7, no. 1, pp. 271-294.

Holmes, N 2015, 'Student perceptions of their learning and engagement in response to the use of a continuous e-assessment in an undergraduate module', *Assessment & Evaluation in Higher Education*, vol. 40, no. 1, pp. 1-14.

Hylton, K, Levy, Y & Dringus, LP 2016, 'Utilizing webcam-based proctoring to deter misconduct in online exams', *Computers & Education*, vol. 92, no. pp. 53-63.

Jordan, AE 2001, 'College student cheating: The role of motivation, perceived norms, attitudes, and knowledge of institutional policy', *Ethics & Behavior*, vol. 11, no. 3, pp. 233-247.

Keeves, JP 1994, *Assessments in Schools: Methods of assessment in schools,* Pergamon Press, Oxford.

Khare, A & Lam, H 2008, 'Assessing student achievement and progress with online examinations: Some pedagogical and technical issues', *International Journal on Elearning*, vol. 7, no. 3, pp. 383-402.

King, CG, Guyette Jr, RW & Piotrowski, C 2009, 'Online exams and cheating: An empirical analysis of business students' views', *Journal of Educators Online*, vol. 6, no. 1.

Knight, P 2001, *A briefing on key concepts: Formative and summative, criterion and norm-referenced assessment,* Learning and Teaching Support Network, York, UK.

Knight, P 2002, 'Summative assessment in higher education: practices in disarray', *Studies in Higher Education*, vol. 27, no. 3, pp. 275-286.

Ladyshewsky, RK 2015, 'Post-graduate student performance in "supervised in-class" vs."unsupervised online" multiple choice tests: implications for cheating and test security', *Assessment & Evaluation in Higher Education*, vol. 40, no. 7, pp. 883-897.

Lancaster, T & Clarke, R 2007, 'The phenomena of contract cheating', in T Roberts (ed.), *Student plagiarism in an online world: Problems and solutions*, Information Science Reference, Hershey, PA, pp. 144-158.

Lancaster, T, Clarke, R & Bretag, T 2016, 'Contract cheating: the outsourcing of assessed student work', in T Bretag (ed.), *Handbook of academic integrity*, Springer, Singapore, pp. 639-654.

Luce, RD & Suppes, P 2002, 'Representational measurement theory'. in J Wixted and H Pashler (eds), *Stevens' handbook of experimental psychology*, Wiley, New York, pp. 1-41.

Maier, U, Wolf, N & Randler, C 2016, 'Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types', *Computers & Education*, vol. 95, pp. 85-98.

Marks, AM & Cronje, JC 2008, 'Randomised items in computer-based tests: Russian roulette in assessment?', *Journal of educational technology & society*, vol. 11, no. 4, pp. 41-50.

McNeill, PM 1993, *The ethics and politics of human experimentation,* CUP Archive, Cambridge, UK.

Messick, S 1995, 'Standards of validity and the validity of standards in performance asessment', *Educational measurement: Issues and practice*, vol. 14, no. 4, pp. 5-8.

Moss, PA 1992, 'Shifting Conceptions of Validity in Educational Measurement: Implications for Performance Assessment', *Review of educational research*, vol. 62, no. 3, pp. 229-258.

Mouberry, SA 2009, 'Motivational predictors of academic cheating among first-year college students: Goals, expectations, and costs', PhD Thesis, North Carolina State University, Raleigh, NC.

Murdock, TB & Anderman, EM 2006, 'Motivational perspectives on student cheating: Toward an integrated model of academic dishonesty', *Educational psychologist*, vol. 41, no. 3, pp. 129-145.
Nahmias, S 1982, 'Perishable inventory theory: A review', *Operations research*, vol. 30, no. 4, pp. 680-708.
Oliver, R 2006, 'Reusable Resources and Authentic Learning Environments', in A Herrington & J Herrington (eds), *Authentic learning environments in higher education*, IGI Global, London, UK, pp. 244-261.
Paas, F, Renkl, A & Sweller, J 2003, 'Cognitive load theory and instructional design: Recent developments', *Educational psychologist*, vol. 38, no. 1, pp. 1-4.
Park, C 2003, 'In other (people's) words: Plagiarism by university students – literature and lessons', *Assessment & evaluation in higher education*, vol. 28, no. 5, pp. 471-488.
Pedhazur, EJ & Schmelkin, LP 2013, *Measurement, design, and analysis: An integrated approach*, Psychology Press, New York.
Pellegrino, JW, Chudowsky, N & Glaser, RE 2001, *Knowing what students know*, National Academies Press, Washington, DC.
Perry, B 2010, 'Exploring academic misconduct: Some insights into student behaviour', *Active Learning in Higher Education*, vol. 11, no. 2, pp. 97-108.
Prince, DJ, Fulton, RA & Garsombke, TW 2009, 'Comparisons of proctored versus non-proctored testing strategies in graduate distance education curriculum', *Journal of College Teaching and Learning*, vol. 6, no. 7, pp. 51.
Ramaprasad, A 1983, 'On the definition of feedback', *Behavioral Science*, vol. 28, no. 1, pp. 4-13.
Ravasco, G 2012, 'Technology-aided cheating in open and distance e-learning', *Asian Journal of Distance Education*, vol. 10, no. 2, pp. 71-77.
Reeves, TC & Hedberg, JG 2007, 'Evaluation strategies for open and distributed learning environments', in BH Khan (ed.), *Flexible Learning in an Information Society*, pp. 226-235.
Reichle, ED, Carpenter, PA & Just, MA 2000, 'The neural bases of strategy and skill in sentence–picture verification', *Cognitive psychology*, vol. 40, no. 4, pp. 261-295.
Rettinger, DA, Jordan, AE & Peschiera, F 2004, 'Evaluating the motivation of other students to cheat: A vignette experiment', *Research in Higher Education*, vol. 45, no. 8, pp. 873-890.
Rigby, D, Burton, M, Balcombe, K, Bateman, I & Mulatu, A 2015, 'Contract cheating & the market in essays', *Journal of Economic Behavior & Organization*, vol. 111, Supplement C, pp. 23-37.
Risko, EF & SJ Gilbert 2016, 'Cognitive offloading', *Trends in cognitive sciences*, vol. 20, no. 9, pp. 676-688.
Rovai, AP 2000, 'Online and traditional assessments: what is the difference?', *The Internet and Higher Education*, vol. 3, no. 3, pp. 141-151.
Russell, MK 2010, 'Technology-aided formative assessment of learning'. in H Andrade and GJ Cizek (ed.), *Handbook of Formative Assessment*, Routledge, Oxford, pp. 125-138.
Sander, P 2016, Using Learning Analytics to Predict Academic Outcomes of First-year Students in Higher Education, master's thesis, University of Oregon, Eugene, OR, viewed at <https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/21969/Sander2016.pdf>.
Schultz, M, Schultz, J & Round, G 2008, 'Online non-proctored testing and its affect on final course grades', *Business Review, Cambridge*, vol. 9, no. pp. 11-16.
Scouller, K 1998, 'The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay', *Higher Education*, vol. 35, no. 4, pp. 453-472.
Sewell, J, Frith, KH & Colvin, MM 2010, 'Online assessment strategies: A primer', *Journal of Online Learning and Teaching*, vol. 6, no. 1, pp. 297-305.

Simon, HA 1982, *Models of bounded rationality: Empirically grounded economic reason*, MIT press, Cambridge, MA.

Stevens, SS 1959, 'Measurement, psychophysics, and utility', in CW Churchman & P Ratoosh (eds), *Measurement: Definitions and theories*, Wiley, New York, pp. 18-63.

Storm, BC, Stone, SM & Benjamin, AS 2017, 'Using the Internet to access information inflates future use of the Internet to access other information', *Memory*, vol. 25, no. 6, pp. 717-723.

Styron, J & Styron Jr, RA 2010, 'Student cheating and alternative web-based assessment', *Journal of College Teaching and Learning*, vol. 7, no. 5, pp. 37.

Suppes, P 2009, 'Measurement Theory and Engineering', in A Meijers (ed.), *Handbook of the Philosophy of Science. Volume 9: Philosophy of Technology and Engineering Sciences*, Elsevier, Burlington, MA, pp. 825-860.

Suppes, P & Zinnes, JL 1962, *Basic measurement theory*, Institute for Mathematical Studies in the Social Sciences, Stanford Univeristy, CA.

Taras, M 2005, 'Assessment – summative and formative – some theoretical reflections', *British journal of educational studies*, vol. 53, no. 4, pp. 466-478.

Tian, X 2007, 'Do assessment methods matter? A sensitivity test', *Assessment & Evaluation in Higher Education*, vol. 32, no. 4, pp. 387-401.

Trenholm, S 2007, 'A review of cheating in fully asynchronous online courses: A math or fact-based course perspective', *Journal of Educational Technology Systems*, vol. 35, no. 3, pp. 281-300.

Trumbull, E & Lash, A 2013, 'Understanding formative assessment: Insights from learning theory and measurement theory', *WestEd*, April.

Ullah, A, Xiao, H & Barker, T 2016, 'A classification of threats to remote online examinations', *Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016 IEEE 7th Annual*, University of British Columbia, Vancouver, 13-15 October, viewed at <http://researchprofiles.herts.ac.uk/portal/files/13603210/Accepted_Manuscript.pdf>.

Underwood, JD 2006, 'Digital technologies and dishonesty in examinations and tests', Qualifications and Curriculum Authority, Coventry, UK.

Whitley, BE 1998, 'Factors associated with cheating among college students: A review', *Research in higher education*, vol. 39, no. 3, pp. 235-274.

Wiliam, D 2011, 'What is assessment for learning?', *Studies in Educational Evaluation*, vol. 37, no. 1, pp. 3-14.

Wiliam, D & Thompson, M 2017, 'Integrating assessment with learning: What will it take to make it work?'. in CA Dwyer (ed.), *The future of assessment: shaping teaching and learning*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 53-82.

Young, JR 2001, 'Texas colleges collaborate to offer online students convenient proctored tests', *Chronicle of Higher Education*, vol. 47, no. 26, pp. A43-A46.

Zanjani, N, Edwards, SL, Nykvist, S & Geva, S 2017, 'The important elements of LMS design that affect user engagement with e-learning tools within LMSs in the higher education sector', *Australasian Journal of Educational Technology*, vol. 33, no. 1, pp. 19-31.