

A model of the pre-assessment learning effects of summative assessment in medical education

Francois J. Cilliers · Lambert W. T. Schuwirth · Nicoline Herman · Hanelie J. Adendorff · Cees P. M. van der Vleuten

Received: 6 August 2010 / Accepted: 18 March 2011 / Published online: 3 April 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract It has become axiomatic that assessment impacts powerfully on student learning. However, surprisingly little research has been published emanating from authentic higher education settings about the nature and mechanism of the pre-assessment learning effects of summative assessment. Less still emanates from health sciences education settings. This study explored the pre-assessment learning effects of summative assessment in theoretical modules by exploring the variables at play in a multifaceted assessment system and the relationships between them. Using a grounded theory strategy, in-depth interviews were conducted with individual medical students and analyzed qualitatively. Respondents' learning was influenced by task demands and system design. Assessment impacted on respondents' cognitive processing activities and metacognitive regulation activities. Individually, our findings confirm findings from other studies in disparate non-medical settings and identify some new factors at play in this setting. Taken together, findings from this study provide, for the first time, some insight into how a whole assessment system influences student learning over time in a medical education setting. The findings from this authentic and complex setting paint a nuanced picture of how intricate and multifaceted interactions between various factors in an assessment system interact to influence student learning. A model linking the sources, mechanism and

F. J. Cilliers (✉) · N. Herman · H. J. Adendorff
Centre for Teaching and Learning, Stellenbosch University, Private Bag X1,
Matieland 7602, South Africa
e-mail: fjc@sun.ac.za

N. Herman
e-mail: nherman@sun.ac.za

H. J. Adendorff
e-mail: hja@sun.ac.za

L. W. T. Schuwirth · C. P. M. van der Vleuten
Department of Educational Development and Research, University of Maastricht,
Universiteitssingel 60, 6229 ER Maastricht, The Netherlands
e-mail: l.schuwirth@maastrichtuniversity.nl

C. P. M. van der Vleuten
e-mail: c.vandervleuten@maastrichtuniversity.nl

consequences of the pre-assessment learning effects of summative assessment is proposed that could help enhance the use of summative assessment as a tool to augment learning.

Keywords Determinants of action · Health sciences education · Higher education · Learning effects of assessment · Medical education · Pre-assessment effects · Mechanism of impact · Model · Source of impact · Summative assessment

Abbreviations

CPA	Cognitive processing activities
MRA	Metacognitive regulation activities
HE	Higher education
HSE	Health sciences education
LESA	Learning effects of summative assessment
SA	Summative assessment

Introduction

Summative assessment (SA) carries inescapable consequences for students and defines a major component of the learning environment's impact on student learning (Becker et al. 1968; Snyder 1971). Consequently, better utilization of assessment to influence learning has long been a goal in higher education (HE), though not one that has been met with great success (Gijbels et al. 2009; Heijne-Penninga et al. 2008; Nijhuis et al. 2005). An extensive search of the literature in various fields suggests that the learning effects of SA (LESA) in authentic HE settings are less well documented than is widely accepted. Literature from HE settings is fragmentary, that from health sciences education (HSE) settings sparse.

Dochy et al. (2007) distinguish pre-, post- and pure learning effects of assessment. Pre-assessment effects impact learning before assessment takes place and are addressed in literature on exam preparation (e.g., van Etten et al. 1997) and test expectancy (e.g., Hakstian 1971). Post-assessment effects impact after assessment and are addressed in literature referring to feedback (e.g., Gibbs and Simpson 2004) and the relationship of assessment with student achievement (e.g., Sundre and Kitsantas 2004). Pure assessment effects impact during assessment and are reported more rarely (Tillema 2001). The testing effect (e.g., Roediger and Butler 2011) could be classified as a pure or a post assessment effect depending on whether the effect on the learning process or subsequent achievement is considered.

Our interest is in the pre-assessment effects of SA on student learning behaviors, specifically in theoretical modules in authentic educational settings. Two major sets of effects can be distinguished: those related to perceived demands of the assessment task and those related to the design of the assessment system.

Perceived task demands

Learning is influenced by students' perceptions of the demands of an assessment task which may accrue from explicit and implicit information from lecturers, from fellow students, past papers and students' own experience of assessment (Entwistle and Entwistle 1991; Frederiksen 1984; van Etten et al. 1997). These perceptions differ from one student to the next (Sambell and McDowell 1998; Scouller 1998; Scouller and Prosser 1994;

Segers et al. 2008, 2006; Tang, 1994). Two types of demands may be distinguished: content demands and processing demands.

Content demands

Content demands relate to the knowledge required to respond to an assessment task (Broekkamp and van Hout-Wolters 2007). These influence what resources students utilize to prepare for assessment by way of cues inferred from the assessor and the assessment task (Entwistle and Entwistle 1991; Frederiksen 1984; Newble and Jaeger 1983; Säljö 1979). They also influence the selection of what content to learn from selected resources. Students cover more content for selected response items than for constructed response items (Sambell and McDowell 1998) and tend to focus on smaller units of information for selected response assessments than for essays (Hakstian 1971, quoting various studies).

Processing demands

Processing demands relate to “skills required for processing ... knowledge in order to generate the requested response” (Broekkamp and van Hout-Wolters 2007). These influence students’ approach to learning by way of cues inferred from the assessor (Ramsden 1979) and from the assessment task. Constructed response items and open-ended assessments are more likely to engender a transformative or deep approach to learning; selected response items and closed assessments, a reproductive or surface approach (Laurillard 1984; Ramsden 1979; Sambell and McDowell 1998; Sambell et al. 1997; Scouller 1998; Tang 1994; Thomas and Bain 1984; van Etten et al. 1997; Watkins 1982). Surprisingly, however, closed-book tests promoted a deep approach to learning more than open-book tests (Heijne-Penninga et al. 2008).

Students who perceived constructed response items to assess higher levels of cognitive functioning were more likely to adopt transformative approaches to learning, even though the questions were adjudged to only assess facts or comprehension (Marton and Säljö 1984). The converse was not true (Scouller and Prosser 1994). Where students perceived essay questions to require simply lifting information from books, learning was not transformative (Sambell et al. 1997). Where students who were intrinsically motivated to understand content perceived an assessment to require memorization, they would memorize facts after having first understood them (Tang 1994).

Tang (1994) speculated that students’ degree of familiarity with an assessment method influenced their approach to learning, Watkins and Hattie (cited by Scouller 1998) that past success with surface strategies may encourage a perception that “deep level learning strategies are not required to satisfy examination requirements” (p. 454).

System design

The mere fact of assessment motivates students to learn and influences the amount of effort expended on learning (van Etten et al. 1997). The amount of time students spend studying increases, up to a point, as the volume of material and, independent of that, the degree of difficulty of the material, to be studied, increases (van Etten et al. 1997). High workloads also drive students to be more selective about what content to study and to adopt low-level cognitive processing tactics (Entwistle and Entwistle 1991; Ramsden 1984; van Etten et al. 1997). The scheduling of assessment in a course and across courses impacts the distribution

of learning effort, as do competing interests e.g., family, friends and extracurricular activities (Becker et al. 1968; Miller and Parlett 1974; Snyder 1971; van Etten et al. 1997).

Theoretical underpinnings

Little previous work on LESA has invoked theory, nor are there many models offering insight into why assessment has the impact it does. Becker et al. (1968) constructed the 'grade point average perspective' to explain the impact of assessment in their setting. Others (Ross et al. 2006, 2003; Sundre and Kitsantas 2004; van Etten et al. 1997) have looked to self-regulation theory. Alderson and Wall (1993) posited that motivation may play a role, but little empirical evidence supported either this contention or a model (Bailey 1996) of washback. Broekkamp and van Hout-Wolters (2007) derived a model from extensive (largely school-based) literature to explain students' adaptation of study strategies when preparing for classroom tests. None of these models has provided a satisfactory explanation of the LESA, however.

The purpose of this study was to extend what is known about the pre-assessment LESA in theoretical modules in a HSE setting by exploring the influence of an authentic, multifaceted assessment system. Specifically, the study aimed to answer three questions: What facets of SA in theoretical modules impact on student learning? What facets of student learning are influenced by such assessment? and In what way do these factors impact on student learning? The intention was to propose a model explaining the pre-assessment LESA of theory using a qualitative approach based on in-depth interviews with medical students.

Methods

This study was conducted at a South African medical school with a six-year, modular curriculum. Phases One and Two comprised three semesters of preclinical theoretical modules; Phase Three, semesters four to nine, alternating clinical theory and clinical practice modules; Phase Four, semesters 10–12, clinical practice modules only. Study guides that included outcomes and information on assessment were provided for each module. Students had to pass each module during the year to access the end-of-year examination and pass each module's examination to progress to the next year of study. Failed modules had to be repeated. While students could repeat multiple modules in a repeat year, they could only repeat 1 year of study. Attrition in the program was approximately 25% overall and was highest during years one and two.

Ethical approval for this study was obtained from an institutional research ethics board. Respondents were informed about the nature of the study, invited to participate and informed consent for study participation and later access to respondents' academic records elicited in writing.

A process theory approach (Maxwell 2004) informed this study. We adopted grounded theory as our research strategy, making a deliberate decision to start with a clean slate and thus utilized in-depth interviews (Charmaz 2006; DiCicco-Bloom and Crabtree 2006; Kvale 1996). This approach offers the advantage of potentially discovering constructs and relationships not previously described. Interviews were not structured beyond exploring three broad themes i.e., how respondents learned, what assessment they had experienced and how they adapted their learning to assessment, all across the entire period of their studies up to that point. Detailed information about the facets of assessment to which respondents adapted their learning and the facets of learning that they adapted in response to assessment were sought throughout, using probing questions where appropriate. When

Table 1 Summary of respondent characteristics based on year of study, gender and academic performance across all 6 years of study

Year of study	Gender	Average mark		
		<70%	70–79%	≥80%
4	F	Resp13		
	M	Resp7 Resp16		
5	F	Resp6	Resp2	Resp4
		Resp12 ^a	Resp11 ^a	Resp8
		Resp15 ^a		Resp9
		Resp17		Resp18
	M	Resp3	Resp5 Resp14	Resp1 Resp10

Resp respondent

^a Respondent failed one/more modules during their studies

new themes emerged in an interview, these were explored in depth. Evidence was sought in subsequent interviews both to confirm and disconfirm the existence and nature of emerging constructs and relationships. In keeping with the grounded theory strategy used, data analysis commenced even as interviews proceeded, with later interviews being informed by preliminary analysis of earlier interviews.

Interviews were conducted with 18 medical students (Table 1) by an educational adviser involved in curriculum reform in the faculty who interacted primarily with lecturers. All students in the fourth (N = 141) and fifth (N = 143) year classes were invited to participate in the study. Each class was addressed once about the study and all students subsequently sent an individual email, inviting them to participate. Thirty-two students volunteered to participate. Interviews were scheduled based on the availability of students, at a time and place of their preference and were conducted in Afrikaans or English based on respondent preference. Notwithstanding the individualized nature of interviews, which each lasted approximately 90 min, no new findings emerged during data collection subsequent to interview 14. After interview 18, the remaining volunteers were thanked but not interviewed.

All interviews were audio recorded, transcribed verbatim and reviewed as a whole, along with field notes. Data analysis was inductive and iterative. Emerging constructs and relationships were constantly compared within and across interviews and refined (Charmaz 2006; Dey 1993; Miles and Huberman 1994). Initial open coding was undertaken by one investigator, subsequent development, revision and refinement of categories and linkages through discussions between the team members. Once the codebook was finalized, focused coding of the entire dataset was undertaken. No new constructs emerged from the analysis of interviews 13–18.

Results

Analysis revealed two sources of impact and two LESA in this setting (Fig. 1). Combining this data with a previously proposed mechanism of impact (Cilliers et al. 2010) allows the construction of the model proposed in the figure.

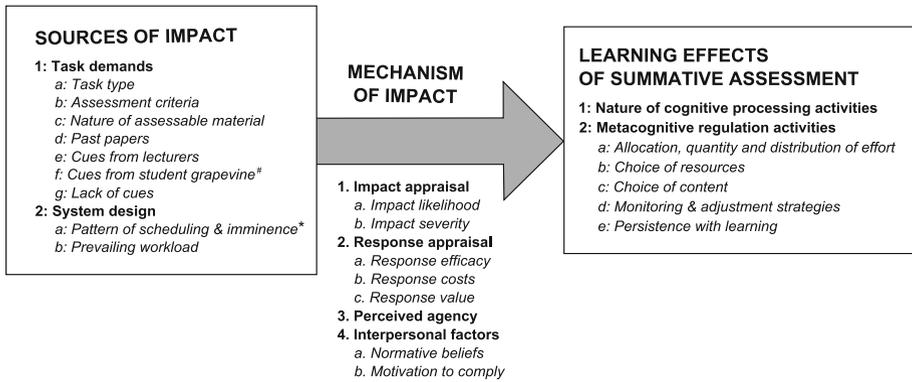


Fig. 1 A model of the pre-assessment LESA. Two sources of impact acted via four facets of the mechanism to bring about two pre-assessment learning effects. *Student grapevine = the informal communication networks between students. *Imminence of assessment = temporal proximity to assessment

The following example illustrates the interplay of factors.

[Quote 1] ... in the earlier years... say first, second and third year you were thinking more along the lines of... I must pass... I must get through this exam. ... But now, when I started hitting last year and this year you start thinking. I'm actually going to be a doctor. It's no good to me using a way to memorize these facts when I'm not going to be able to use it practically in my job one day. ... I've started trying to change the way I study... the way I approach a module is trying to maybe within the first two or three weeks of a four week block, to understand... to understand the concepts more and to sit with the concepts and try to work the concepts out rather than just memorizing. And then it's unavoidable when it comes to the last week, last week and a half of a block ... You just try and cram—try and get as many of those facts into your head just that you can pass the exam and it involves... sadly it involves very little understanding because when they come to the test, when they come to the exam, they're not testing your understanding of the concept. They're testing "can you recall ten facts in this way?" ... And yes you know that I'm gonna be asked to recall five facts. So then you just learn five facts rather than trying to understand the core concepts. **(Resp1)**

As assessment becomes more *imminent* ("when it comes to the last week, last week and a half of a block"), *impact likelihood* ("it's unavoidable") and *impact severity* ("just that you can pass the exam") are considered, along with *response value* (success in assessment increasing, patient care decreasing in value as assessment looms). These factors, together with *task type* ("they're not testing your understanding of the concept. They're testing 'can you recall ten facts in this way?'") and *response efficacy* ("You just try and cram—try and get as many of those facts into your head just that you can pass the exam") considerations, generate an impact on the *nature of cognitive processing activities (CPA)* ("So then you just learn five facts rather than trying to understand the core concepts").

The interplay between these various factors can be visualized as a three-dimensional matrix. In different scenarios, different sources and mechanisms combined in varying intensities to yield different sets of effects. Not all factors were active in all respondents at

Table 2 Linking the sources of impact and LESA

SOURCES OF IMPACT		LEARNING EFFECTS OF SUMMATIVE ASSESSMENT				
		EF1: Nature of Cognitive Processing Activities	EF2: Metacognitive Regulation Activities			
			EF2a: Allocation, quantity and distribution of effort	EF2b: Choice of resources	EF2c: Choice of content	EF2d: Monitoring & adjustment strategies
SF1. Task demands	SF1a: Task type					
	SF1b: Assessment criteria					
	SF1c: Nature of assessable material					
	SF1d: Past papers					
	SF1e: Cues from lecturers					
	SF1f: Cues from student grapevine					
	SF1g: Lack of cues					
SF2. System Design	SF2a: Pattern of scheduling & imminence					
	SF2b: Prevailing workload					

This table provides a composite summary of the interrelationships between source and effect factors for the assessment system as a whole. Where a source factor (SF; row headers) and an effect factor (EF; column headers) were found to be related during data analysis, the intersecting cell in the table has been shaded. All source factors influenced multiple effect factors e.g., past papers influenced CPA, choice of resources, choice of content and monitoring and adjustment strategies. Each effect factor was influenced by several source factors e.g., in addition to being influenced by past papers, monitoring and adjustment strategies were also influenced by task type, the student grapevine and system design

all times or, indeed, the same respondent at different times. Furthermore, different sets of interactions were found to be at play for different respondents in the same assessment context and for the same respondent in different assessment contexts.

The relationship between source and effect factors was mutually multiplex (Table 2) i.e., any given source of impact influenced various learning effects and any given learning effect was influenced by various sources of impact. To illustrate these interrelationships, the two source factors with the greatest scope of influence i.e., pattern of scheduling & imminence, and prevailing workload, will be described, as will the three effects influenced by the greatest number of source factors i.e., nature of CPA, choice of resources and choice of content. For the sake of brevity, only the more prominent features of each will be described.

Pattern of scheduling and imminence (row SF2a, Table 2)

Imminence—the temporal proximity—of assessment strongly influenced respondents’ learning behavior, in a pattern determined by the pattern of scheduling of assessment.

CPA Respondents adopted higher-order CPA when assessment was more distant, lower-order CPA as assessment became more imminent (cf. Quote 1).

Effort While the pattern of scheduling of assessment had the beneficial effect of ensuring that respondents regularly allocated effort to learning, they adopted a periodic rather than a continuous pattern of study. In an effort to devote attention to other aspects of their lives, respondents devoted little or no effort to learning at the start of each module. Interests and imperatives other than learning were relegated to a back seat as assessment loomed, however, and learning effort escalated dramatically.

[Quote 2] I think that because we write tests periodically, I swot* periodically. ... I think [the block method] has definitely caused me to change from a person who worked continuously to a person who does the minimum for the first three weeks of every month and the maximum in the last week. **(Resp2)**

* swot: study, especially intensively in preparation for an examination

Resources Concurrently, though, as assessment became more imminent, the range of resources respondents utilized shrank.

Content Cue-seeking behavior and responsiveness to cues both typically intensified as assessment grew more imminent.

[Quote 3] I ask myself a question, cover up the answers and write it out, see how many I can get. If I only get like five or six then I look through the stuff again. And I close it again, try it again ... if there's time to do that thoroughly with most of the things in my summary, I do, but if I'm running all out of time, then I look at the things that were specifically indicated to me as important and I cram those. **(Resp4)**

Even respondents who evidenced intrinsic motivation and a mastery goal orientation sought out and acted on cues to inform what content to learn.

Persistence While regular, periodic assessment lead to exhaustion, imminent assessment helped motivate respondents to persist in allocating time and effort to learning despite growing fatigue.

Prevailing workload (row SF2b, Table 2)

Respondents perceived the workload in many modules to be beyond their capacity to deal with, a situation sometimes compounded by poor planning by respondents.

CPA Where workload was manageable, higher-order CPA were adopted. Where workload was unmanageable, even respondents who preferred adopting higher-order CPA would utilize lower-order CPA.

[Quote 4] memorizing stuff ... it misses the point, because if you are going to swot like that just to pass a test, what will that help? ... But I mean, sometimes your time is just too little and you must just go and swot the stuff parrot-fashion, but I don't like doing it like that at all. **(Resp10)**

Effort The higher the prevailing workload, the greater the likelihood that effort would be allocated to studies rather than other aspects of respondents' lives. More effort was also expended, distributed more evenly across the duration of the module.

Resources A high workload inhibited the sourcing and utilization of resources other than those provided by lecturers. Only where resources provided by lecturers were considered inadequate did respondents source and utilize other resources, workload notwithstanding.

Content Where workload was manageable, respondents studied content they considered relevant and material promoting understanding and clinical reasoning. Where workload was unmanageable, respondents focused on material more likely to ensure success in assessment, even if this selection conflicted with what they would have learned to satisfy longer-term clinical practice goals.

[Quote 5] ... [names module], it's a massive chunk of work ... you must swot so selectively about what you are going to leave out and it's not as though you leave out less important things. You leave things out that you think they will not ask. So it's maybe big things or maybe important things that could save a patient's life one day, but you don't swot it because you have to pass the test now and that's a problem for me. **(Resp6)**

Monitoring and adjustment While it ensured that respondents devoted appropriate amounts of effort to studying, a high workload could be accompanied by a disorganized rather than systematic approach to MRA.

Nature of CPA (column EF1, Table 2)

Task type Respondents inferred processing demands directly from the item type to be used or indirectly based on the complexity of the cognitive challenge posed (cf. Quotes 1, 11) and adjusted their CPA accordingly.

[Quote 6] Look, medicine is such that there isn't actually much to understand, you just have to know. ... Sometimes there is an application question or so but in general I'd say it's just ... focused on your ability to remember, to recall what you do in the notes. ... there are a few little things that you must understand, like the Physiology ... but as you progress it's mostly just "name two causes". **(Resp9)**

Respondents varied in how they gauged the demands of assessment tasks, however. Some respondents dismissed true–false questions used in Phase 2 of the program as simply requiring recognition of previously encountered facts and responded with low-level CPA. In contrast, other respondents identified these same questions as assessing at a higher cognitive level than any other assessments in the program. They adapted their CPA to achieve understanding of the material so as to be able to reason through the questions. A similar dichotomy was evident in responses to long answer questions.

Assessment criteria Where respondents perceived marking to be inflexibly done according to a predetermined memorandum, they responded with rote memorization to try and ensure exact reproduction of responses.

Nature of assessable material Where material was perceived to be understandable and logical, respondents adopted higher-order CPA. Where material was less understandable or where the level of detail required to understand the logic was too deep, respondents adopted superficial CPA.

[Quote 7] ... [names subject] is not very logical. You cannot reason it out for yourself ... unless you go and look in super-depth ... I have neither the time nor the interest to go and do that. **(Resp10)**

Lecturers Lecturing using PowerPoint to present lists of facts rather than in a manner that helped respondents develop their understanding of a topic cued memorization as a learning response.

[Quote 8] ... if the lecture goes on with a guy that puts up PowerPoint's that click, click, click, click, click and here comes a bunch of information, the next slide. You know, he's not really going to test your insights, because he didn't try and explain the concepts to you at all. He just simply gave you facts. So you can just expect that the paper will be factual. **(Resp10)**

Student grapevine Peers identified certain modules as making higher-order cognitive demands, others as requiring only extensive memorization of material. Respondents geared their CPA accordingly.

The influence of system design on CPA has been detailed earlier.

Choice of resources (column EF2b, Table 2)

Task type Assessment incorporating small projects resulted in respondents sourcing and utilizing resources they would not otherwise have used e.g., textbooks in the library, the internet generally and literature databases more specifically. However, apart from promoting the use of past papers as a resource, most other assessment tasks cued the utilization of less, rather than more, diverse resources.

Past papers The more any given lecturer utilized a particular question type or repeated questions from one assessment event to another, the more respondents utilized past papers to plan their learning and select material to learn.

Lecturers The resources lecturers provided or utilized were perceived to delineate what content was more likely to feature in assessment. Much planning effort was devoted to obtaining copies of PowerPoint slides used, or handouts provided, by lecturers. Some lecturers were perceived as being tied to a particular resource e.g., a prescribed textbook, which respondents then focused on. Equally, use was often not made of textbooks as other resources were perceived to be more appropriate for assessment purposes.

[Quote 9] ... with [names module] that we did last year. They kind of said to us we don't want you to study from the lecture notes we want you to study from [the] prescribed textbook. ... but if we had studied from [names textbook], we wouldn't have passed the test. So, you know, they've got their questions in their notes and that's basically what you have to study to get through. **(Resp1)**

Student grapevine Cues obtained ahead of or early in the course of a module about the likely content of assessment influenced the resources respondents opted to use in preparation for assessment (cf. Quote 13).

The influence of system design on choice of resources has been detailed earlier.

Choice of content (column EF2c, Table 2)

Having selected what resources to utilize, respondents made a second set of decisions about what content to study from those resources. Selecting material to study and selecting material to omit were distinct decisions.

Task type Respondents sought out material they perceived could be asked using any given task type and omitted information they perceived could not (cf. also Quotes 1, 11). Information about the overall extent of assessment, the number of marks devoted to each section of the work and the magnitude of questions also influenced choice of content. For example, if respondents knew there would be no question longer than 10 marks in an assessment, they omitted tracts of work they perceived could only be part of a longer question.

[Quote 10] then like in Patho[logy] or take now Physiology ... you know they cannot ask it, because there's too few marks to ask such a big piece and it's difficult to ask it, so you leave it out. ... Yes, say a question paper can count 200 marks at the most. Then an 80 mark one is not even so much. ... you know they can't just ask 20 marks just about one disease's pathophysiology. They are actually going to try and cover stuff as widely as possible ... So, at the end of the day, you leave pretty important stuff out for now to learn ridiculous lists of things. ... You know you must get these little lists in your head, but at the end of the day, your insight in... in the whole story is... left out. Now if you know the pathophysiology, you can figure out almost anything. ... [I learn lists] because that's what they ask. ... You know then they will say, name five causes of this for me. Or list 10 differential diagnoses of this. You have to give lists. **(Resp6)**

Past papers were used to determine not only what topics but also what kind of material to study or omit (cf. also Quote 13).

[Quote 11] what I will do is I will see a question paper, I browse through the whole paper, the whole paper is just five to ten point "listing" questions. They're not paragraph or insight or case-studies... Or the case studies are actually just again just to get the little lists out of you in another way... And so I will concentrate on that. If I page through that work and here is some or other description of a thing, I'll maybe skip over it, but you'll stop at a five or ten point list. Yes, then one just swots like that, you know, that is what will get you through. **(Resp7)**

Lecturers Direct cues from lecturers included general comments in class like "this will (or won't) be in the exam" and specific "spots" provided to students. Respondents attended to such cues even if they perceived the content identified as important to be irrelevant to later clinical practice.

[Quote 12] the lecturers do rather have a tendency to give spots and then I always feel, okay, I now know that question is going to get asked. I think it is a stupid question to ask, but I suppose it is going to get asked. Then I just go and swot that list before the time **(Resp10)**

Indirect cues accrued from the amount of time devoted to a topic in a module or even a particular lecture. This influenced how much time was devoted to that topic while studying.

Student grapevine Guidance about assessment from senior students and peers influenced respondents' choice of content, even if they considered the material covered by the cues to be irrelevant to their longer-term goal of becoming a good clinician.

[Quote 13] I will easily go and look at an old question paper or two or so, tips that other students give and based on that, I will go... go learn, focus on certain things. And, to my own detriment for the day that clinical comes, skip some things ... I will talk to guys that are a year or two years ahead of me, especially guys that are a year

ahead of me, because the course is still much the same as what they had ... if someone says “listen here, for that little test, they took that book, but at the back of the book there’s a bunch of questions and they really just concentrated on the questions”... Then I’ll maybe read the book too, but I will concentrate on the questions. **(Resp7)**

Lack of cues Where respondents could not discern cues about what to expect in assessment, they typically tried to learn their work more comprehensively, but at the cost of increased anxiety.

The influence of system design on choice of content has been detailed earlier.

Disconfirming evidence and negative cases

No disconfirming evidence or negative cases became apparent during analysis. In all instances, the described sources and effects were a fundamental part of how respondents described their learning response to assessment. There were limited instances where respondents’ learning was not influenced by task demands, even though they perceived different assessment tasks to pose different demands. This was the exception rather than the rule, however, and other dimensions of the responses described above were discernable for these respondents. Patient care goals vied strongly with assessment for prominence in respondents’ academic goal structures (cf. Quotes 1, 5, 10, 12, 13), particularly for more senior respondents. However, as is evident, assessment ultimately trumped patient care in influencing learning.

Discussion

The purpose of this study was to extend what is known about the pre-assessment LESA in theoretical modules in a HSE setting by exploring the influence of an authentic, multi-faceted assessment system. Many of our findings support findings about the LESA emanating from other HE settings, but this is the first time that many of these factors have been shown to be at play in a HSE setting. Our findings also paint a nuanced picture how factors relate to one another and how their influence is modulated. The influence on learning of the imminence of assessment, a newly described factor, is a case in point. Any work investigating the impact of assessment on learning needs to take cognizance of how near or far assessment looms, as responses are likely to vary based on this.

The findings in this paper derive from respondents’ experience of assessment over a number of years, albeit collected at one time point. The influence of assessment has been ascertained in the context of other academic and non-academic influences in a “whole curriculum” setting. This work thus adds to a relatively limited body of literature, particularly for HSE, on the LESA emanating from authentic rather than controlled settings.

Together with other findings (Cilliers et al. 2010) it has been possible to propose a theoretical model not only describing what the pre-assessment LESA are, but also explaining why students interact in the way that they do with assessment (Fig. 1). Self-regulation theory has previously been invoked when discussing the link between assessment and learning (Ross et al. 2006, 2003; Sundre and Kitsantas 2004; van Etten et al. 1997). Our findings suggest that self-regulation does indeed play a role, but that it is part of a broader framework. Our findings also lend empirical support to some aspects of the model proposed by Broekkamp and van Hout-Wolters (2007).

Our model makes no claim for a solo role for SA as a determinant of learning behavior. Instead, it emphasizes how SA influences learning behavior prior to an assessment event and fleshes out the line linking assessment and learning in other models (e.g., Biggs 1978 fig. 1, p. 267; Ramsden 2003 fig. 5.1, p. 82). However, given the profound consequences associated with SA, there can be little doubt that the factors described in this report play a significant role in the overall picture.

While no single assessment-based intervention is likely to affect all students' learning in the same way or to the same degree, what our model does offer is a means of understanding how and why SA impacts on student learning. On the one hand, this should allow for the deliberate use of SA in the design of a stimulating learning environment. While every situation will need to be analyzed on its own merits, our results do suggest some broad focus areas. If the intention is to influence students' CPA, attention should be paid to the assessment methods, the level of cognitive challenge posed and students' perceptions of these. If the intention is to influence the content to which students attend, then various cues in the system need to be managed and aligned. Attending to the prevailing workload and considering the influence of imminence can influence a wide range of MRA. On the other hand, the model can also inform "diagnostic" review of assessment programs where assessment has undesirable effects on student learning.

Along with other reports of attempts to influence learning using assessment (e.g., Gijbels et al. 2009; Heijne-Penninga et al. 2008), we see our study as a cautionary tale to those who would wield assessment to influence learning. In the curriculum that preceded the one upon which this report is based, students studied multiple theoretical modules concurrently and wrote tests on multiple modules at four pre-determined times during the year. This resulted in what was considered an undesirable pattern of learning i.e., little learning effort for 2 or 3 months, followed by binge-learning for a couple of weeks prior to the tests. The modular design of the present curriculum was in part an attempt to induce more continuous and effective learning. However, while students do allocate time to learning more frequently than in the past, the impact appears to have been simply the induction of shorter cycles of binge-learning than had characterized the previous curriculum.

It could be argued that a severe limitation of this work is the fact that the findings are derived qualitatively and from only medical students in one South African setting. However, our findings resonate with and indeed, further illuminate, findings reported elsewhere (Becker et al. 1968; Miller and Parlett 1974; Snyder 1971; van Etten et al. 1997). These reports derived from a range of universities in the United States and Scotland and drew on students from a range of disciplinary backgrounds. This suggests the constructs and relationships described in this report exhibit a certain robustness across a variety of contexts.

Our study population contained a large proportion of successful students. For our qualitative study this can be considered the most sensitive population as these were the students who were most successful in understanding how the assessment game works. As to the generalizability of our findings, this needs to be studied further in more quantitative orientated studies. Other directions for future research include further validating the model and subjecting the many individual relationships within the model to more detailed scrutiny.

The focus when thinking about the impact of assessment on students' learning is often on formative assessment and the role of feedback. However, this study shows how overwhelming the impact of SA on learning can be. No matter how carefully the rest of the

assessment program is designed, given the stakes of SA, the latter will quite likely overwhelm other aspects of assessment if not designed in harmony with the overall program.

Acknowledgments We are grateful to the students for their participation in this project. We would like to thank Dr Susan van Schalkwyk for her insightful comments on the manuscript. This material is based upon work financially supported by the National Research Foundation in South Africa. Financial support from the Foundation for the Advancement of International Medical Education and Research is also gratefully acknowledged.

Conflict of interest No conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, *14*, 115–129.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, *13*, 257–279.
- Becker, H. S., Geer, B., & Hughes, E. C. (1968). *Making the grade: The academic side of college life*. New York: John Wiley & Sons, Inc.
- Biggs, J. B. (1978). Individual and group differences in study processes. *British Journal of Educational Psychology*, *48*, 266–279.
- Broekkamp, H., & van Hout-Wolters, B. H. A. M. (2007). Students' adaptation of study strategies when preparing for classroom tests. *Educational Psychology Review*, *19*, 401–428.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London: Sage Publications Ltd.
- Cilliers, F. J., Schuwirth, L. W. T., Adendorff, H. J., Herman, N., & van der Vleuten, C. P. M. (2010). The mechanism of impact of summative assessment on medical students' learning. *Advances in Health Sciences Education*, *15*, 695–715.
- Dey, I. (1993). What is qualitative analysis? In *Qualitative data analysis* (pp. 31–54). USA and Canada: Routledge.
- DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, *40*, 314–321.
- Dochy, F., Segers, M., Segers, M., Gijbels, D., & Struyven, K. (2007). Assessment engineering: Breaking down barriers between teaching and learning, and assessment. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 87–100). Oxford: Routledge.
- Entwistle, N. J., & Entwistle, A. C. (1991). Contrasting forms of understanding for degree examinations: The student experience and its implications. *Higher Education*, *22*, 205–227.
- Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist*, *39*, 193–202.
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, *2004–05*, 3–31.
- Gijbels, D., Coertjens, L., Vanthourhout, G., Struyf, E., & Van Petegem, P. (2009). Changing students' approaches to learning: A two-year study within a university teacher training course. *Educational Studies*, *35*, 503–513.
- Hakstian, A. R. (1971). The Effects of type of examination anticipated on test preparation and performance. *Journal of Educational Research*, *64*, 319–324.
- Heijne-Penninga, M., Kuks, J. B. M., Hofman, W. H. A., & Cohen-Schotanus, J. (2008). Influence of open- and closed-book tests on medical students' learning approaches. *Medical Education*, *42*, 967–974.
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. (Thousand Oaks, California: Sage Publications).
- Laurillard, D. M. (1984). Learning from problem solving. In F. Marton, et al. (Eds.), *The Experience of learning* (pp. 124–143). Edinburgh: Scottish Academic Press.

- Marton, F., & Säljö, R. (1984). Approaches to learning. In F. Marton, et al. (Eds.), *The experience of learning* (pp. 36–55). Edinburgh: Scottish Academic Press.
- Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, 33, 3–11.
- Miles, M. B. and Huberman, A. M. (1994). Making good sense—Drawing and verifying conclusions. In *Qualitative data analysis: An expanded sourcebook* (pp. 245–287). Thousand Oaks: Sage Publications.
- Miller, C. M. L., & Parlett, M. (1974). *Up to the mark: A study of the examination game*. London: Society for Research into Higher Education.
- Newble, D. I., & Jaeger, K. (1983). The effect of assessment and examinations on the learning of medical students. *Medical Education*, 17, 165–171.
- Nijhuis, J. F. H., Segers, M. S. R., & Gijsselaers, W. H. (2005). Influence of redesigning a learning environment on student perceptions and learning strategies. *Learning Environments Research*, 8, 67–93.
- Ramsden, P. (1979). Student learning and perceptions of the academic environment. *Higher Education*, 8, 411–427.
- Ramsden, P. (1984). The context of learning. In F. Marton, et al. (Eds.), *The experience of learning* (pp. 144–164). Edinburgh: Scottish Academic Press.
- Ramsden, P. (2003). *Learning to teach in higher education*. London: Routledge.
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15, 20–27.
- Ross, M. E., Green, S., Salisbury-Glennon, J. D., & Tollefson, N. (2006). College students' study strategies as a function of testing: An investigation into metacognitive self-regulation. *Innovative Higher Education*, 30, 361–375.
- Ross, M. E., Salisbury-Glennon, J. D., Guarino, A., Reed, C. J., & Marshall, M. (2003). Situated self-regulation: Modeling the interrelationships among instruction, assessment, learning strategies and academic performance. *Educational Research and Evaluation*, 9, 189–209.
- Säljö, R. (1979). Learning about learning. *Higher Education*, 8, 443–451.
- Sambell, K., & McDowell, L. (1998). The construction of the hidden curriculum: Messages and meanings in the assessment of student learning. *Assessment and Evaluation in Higher Education*, 23, 391–402.
- Sambell, K., McDowell, L., & Brown, S. (1997). 'But is it fair?': An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23, 349–371.
- Scouller, K. M. (1998). The influence of assessment method on students' learning approaches: Multiple choice examinations versus assignment essay. *Higher Education*, 35, 453–472.
- Scouller, K. M., & Prosser, M. (1994). Students' experiences of studying for multiple choice question examinations. *Studies in Higher Education*, 19, 267–279.
- Segers, M., Gijbels, D., & Thurlings, M. (2008). The relationship between students' perceptions of portfolio assessment practice and their approaches to learning. *Educational Studies*, 34, 35–44.
- Segers, M., Nijhuis, J., & Gijsselaers, W. (2006). Redesigning a learning and assessment environment: The influence on students' perceptions of assessment demands and their learning strategies. *Studies in Educational Evaluation*, 32, 223–242.
- Snyder, B. R. (1971). *The Hidden Curriculum*. New York: Alfred A. Knopf.
- Sundre, D. L., & Kitsantas, A. (2004). An exploration of the psychology of the examinee: Can examine self-regulation and test-taking motivation predict consequential and non-consequential test performance? *Contemporary Educational Psychology*, 29, 6–26.
- Tang, K. C. C. (1994). Effects of modes of assessment on students' preparation strategies. In G. Gibbs (Ed.), *Improving student learning—theory and practice* (pp. 151–170). Oxford: Oxford Centre for Staff Development, Oxford Brookes University.
- Thomas, P. R., & Bain, J. D. (1984). Contextual dependence of learning approaches: The effects of assessments. *Human Learning*, 3, 227–240.
- Tillema, H. H. (2001). Portfolios as developmental assessment tools. *International Journal of Training and Development*, 5, 126–135.
- van Etten, S., Freebern, G., & Pressley, M. (1997). College students' beliefs about exam preparation. *Contemporary Educational Psychology*, 22, 192–212.
- Watkins, D. (1982). Factors influencing the study methods of Australian tertiary students. *Higher Education*, 11, 369–380.