The hidden curriculum revisited: a critical review of research into the influence of summative assessment on learning

Gordon Joughin*

Academic Development Unit, University of Wollongong, Wollongong, Northfields Avenue, NSW 2522, Australia

That summative assessment drives learning has become one of the most frequently stated maxims in the literature of assessment and learning in higher education. However, a careful review of the empirical research often cited in support of this proposition may cause us to reconsider its veracity and to seek a more nuanced understanding of the research findings. Seminal research on the influence of assessment on learning in higher education was reviewed on the basis of its context, research methods used, sampling, reported findings, and the generalisability of these findings. This study found that the treatment of the research reported in frequently cited works such as Making the grade, The hidden curriculum, and Up to the mark has often oversimplified, and thus misrepresented, the research findings, leading to singular interpretations of complex, multi-faceted phenomena. Other research suggesting limitations to the capacity of assessment per se to improve students’ approaches to learning is often misunderstood or under-emphasised, leading to the risk of exaggerated claims for the capacity of ‘alternative’ forms of assessment to foster effective learning processes in students. The findings of this review lead to a proposed agenda for empirical research to address what seem to be serious gaps in our understanding of fundamental aspects of the interactions between assessment and learning.

Keywords: hidden curriculum; approaches to learning; critical review; assessment and learning; cue-consciousness

Introduction

No academical exercise can be more useful than that of examination. It whets the desire for learning; it enhances the solicitude of study while it animates the attention of whatever is taught. (Philip Melancthon, De Studs Adolescentum, fifteenth century, cited in Madaus, Russell, and Higgins 2009, 140)

The belief that summative assessment drives students’ learning is widespread, long-standing, and deeply held in higher education circles. Students’ experience of summative assessment has been the focus of many important and rightly influential studies that confirm the central role played by assessment in students’ academic lives. However, the empirical basis of this belief is rarely articulated, with foundational research cited frequently but rarely elaborated upon or considered critically. This short paper seeks to review the role of summative assessment in student learning primarily through a re-consideration of three pivotal, sociologically oriented
studies on the role of assessment in students’ academic lives, supplemented by a short look at some representative studies of the differential influence of assessment methods on students’ study. It will be argued that these studies are commonly misunderstood, that generalisations drawn from them are often not warranted, and that new research is required to determine the extent to which, and in what ways, students’ experience of assessment influences their patterns of and approaches to study.

Three classic studies, published within six years of each other, are at the heart of this exploration: Becker, Geer, and Hughes’ *Making the grade: The academic side of college life* (1968); Snyder’s *The hidden curriculum* (1971); and Miller and Parlett’s *Up to the mark: A study of the examination game* (1974). In the literature of higher education, these works are ubiquitous. In a recent review article, Stevens, Armstrong, and Arum (2008) note that ‘(s)ome of the most insightful sociological investigations of college life are more than 30 years old’, naming Becker et al.’s work amongst these (131) though the other two could equally have been included. These works were published at a time when sociological studies of organisational behaviour were being particularly well received in academic and popular circles. The research underpinning them was well funded, the works were written in highly accessible styles designed for academic but non-specialist audiences, and their findings were expressed clearly, sometimes dramatically, and in ways that led the reader into the life world of contemporary students. Their findings quickly became part of the intellectual discourse on summative assessment in higher education, as well as entering into the folklore of university teachers on several continents, and their influence has continued to the present day.

In this paper:

- the widespread use of these works will be noted;
- each study will be summarised according to its context, methodology, sample, findings, and generalisability;
- related research in the ‘student approaches to learning’ tradition will be noted; and
- a new research agenda to explore the contemporary student experience of assessment will be proposed.

The classic studies

The three works noted above have been widely cited since their respective publications, and their central ideas have been acknowledged in numerous influential books, articles, papers, and reports.

Most key texts on assessment in higher education published in the UK make reference to one or more of these works. Entwistle’s opening chapter to *The experience of learning* (1997) includes references to all three works. Ramsden (2003), in his well known and much cited *Learning to teach in higher education*, begins his discussion of the effects of assessment on learning with supportive references to the works of both Snyder and Becker et al., using Snyder’s study of the ‘hidden curriculum’ to support the primacy of assessment in students’ perceptions and reinforcing this with explicit reference to Becker et al. (67). Gibbs and his colleagues similarly draw on all three works in their research on assessment and learning (Dunbar-Goddet et al. 2006; Gibbs 2006a, 2006b). Sambell and McDowell (1998) firmly locate their research in the context of the ‘hidden curriculum’ as represented by Snyder and Becker et al.
Rowntree (1987) adopts the term ‘hidden curriculum’ from Snyder, provides a brief summary of Miller and Parlett’s research, and includes several references to Becker et al. Bloxham and Boyd (2007) cite Snyder alone to support their contention that ‘the assessment strategy of a particular course has a major impact on student activity’ (16) and turn to Miller and Parlett in claiming that students become ‘cue conscious’ (19).

Elliott (2008, 282) cites all three works in the space of two sentences:

Studies conducted by Becker et al. (1968), Entwistle and Wilson (1977), Miller and Parlett (1974) and Snyder (1971) appeared to confirm the powerful influence of assessment. Miller and Parlett (1974), for example, identified different degrees of ‘cue-consciousness’, that is ‘the extent to which students recognized or actively sought out “cues” from staff to help guide their attempts to pay the assessment game’ (Entwistle 1984, 16).

The above examples are not isolated instances; these three works have been frequently and consistently cited since their publication. For example, a ‘cited reference’ search of the leading journals in the ISI Web of Knowledge database listed 223 citations of Making the grade, spread more or less evenly across the past three decades, and 142 citations of The hidden curriculum, though admittedly only 11 for Up to the mark. This database is restricted to high impact journals and does not include some important higher educational journals such as Assessment & Evaluation in Higher Education or Higher Education Research & Development. A perhaps cruder but more comprehensive search using Google Scholar identified 22 articles citing all three authors.

After at least 35 years, these works continue to be cited as frequently as ever. The following, again, are merely some recent examples of articles referring to one or more of these works: Orrell (2006), Hawe (2007), Kvale (2007), Dochy et al. (2007), Nespor (2007), Titus (2008), and Etten et al. (2008).

Given the reliance that is frequently placed upon them, it is particularly important that we have a clear understanding of what these works have to say about assessment and learning, the nature of the studies they report, and the applicability of their findings to contemporary contexts. The following summaries present the contexts, methods, findings, and limitations of each in the order of their publication.

**Making the grade (1968)**

*Authors.* Becker, Geer, and Hughes were sociologists with well-established reputations and had previously authored an influential study of medical education.

*Context.* Making the grade is based on a study conducted at the University of Kansas between 1959 and 1961. The university had 7004 residential undergraduates across seven professional schools and a liberal arts school.

*Sample.* The sample is not specified; the approach can involve relating to a wide cross-section of students, and apparently all schools were involved.

*Method.* The study is based on symbolic interactionism and participant observation, though the authors do not appear to have taken on student roles. This approach focuses on collective action carried out in a context of shared meaning. The concept of ‘perspective’ is used to analyse these actions, which are seen to be the ‘collective actions of relatively homogeneous groups’ (5).
Finding. The overarching finding was that students, as a group, adopted what the authors’ termed ‘the GPA (Grade Point Average) perspective’ – grades were seen as the major institutionalised valuable of the campus, with only a small minority of students valuing learning for its own sake.

Limitations. Since the research method assumed ‘collective actions of a relatively homogenous group’ (5), the study was unlikely to identify variation in students’ responses to their environment. By virtue of the research framework, it was assumed that the student group was homogenous, and that they responded collectively. We would not now make such assumptions, given the variation that is known to occur in most aspects of students’ experience of learning.

Generalisability. The authors are adamant that the GPA perspective represents an institutionalisation process characteristic of the particular university studied and that they would not expect all campuses to be like it. Moreover, they state that ‘on a campus where something other than grades was the major institutionalized form of value, we would not expect the GPA perspective to exist at all’ (122):

Let us be clear … about what we mean. We do not mean that all students invariably undertake the actions we have described, or that students never have any other motive than getting good grades. We particularly do not mean that the perspective is the only possible way for students to deal with the academic side of campus life. (121)

The authors clearly believe that generalisations from their study need to be made very cautiously.

The hidden curriculum (1971)

Author. Snyder was a psychiatrist and dean of Institute Relations at the Massachusetts Institute of Technology (MIT) at the time of his study.

Context. This study was undertaken at the MIT between 1961 and 1964. MIT operated on ‘severe selectivity’ so students were highly able and highly motivated, with a demanding curriculum, work overload and apparent reward for ‘measured achievement’. The author states that he also considered students at Wellesley College, a small (2000 students) elite (top 5% of students) institution for women, though he does not refer to any of its students in his report.

Sample. The study focussed on first-year engineering and science students. The author reports that a ‘random’ sample of 54 students were interviewed – no details of the sample are provided, and data derived from these interviews are not presented in any systematic way. Rather, five students were selected to represent the range of responses. Four of these students were outstandingly successful in their studies, while one appeared to have dropped out. No information is provided regarding Wellesley students and no separate data are reported on them.

Method. No details of the research method are provided, other than it involved (1) interviews of the MIT students, and (2) a reliance on the author’s 10-year-old memories of students at Wellesley College.
Findings. This work is well known for popularising the term ‘the hidden curriculum’ in higher education circles: while assessment tasks are stated explicitly, there is a parallel set of expectations based on the social context and the assumptions and expectations of both teachers and students. Thus, ‘(e)ach student figures out what is actually expected as opposed to what is formally required’ (9) and adapts their studying accordingly.

Limitations. The sample is dominated by male engineering and science students at an elite institution characterised by an expectation of high achievement. The research method is not specified, it relies in part on 10-year-old memory, and, somewhat extraordinarily, the only data presented are drawn from a sample of five students.

Generalisability. While the notion of the ‘hidden curriculum’ has resonated strongly with higher education teachers, it is difficult to justify any generalisation of this study’s findings to any wider population of students, given the elite nature of the sample and context, and the narrow and small sample reported on.

Up to the mark: A study of the examination game (1974)

Authors. Miller and Parlett were respected researchers – one an experimental psychologist and one an educational researcher.

Context. The study was undertaken at the University of Edinburgh. The relevant part of the study involved honours students preparing for their final examination in law. This was a small class (34) with close contact between staff and students, within an elite and highly selective university. The overall study covered four aspects of assessment. The relevant section of the report, on cue-consciousness and examinations, constitutes 22 of its 128 pages.

Sample. The sample consisted of 30 final-year law students.

Method. The authors used an ‘illuminative’ approach, ‘broadly in the symbolic interactionist tradition’ and influenced by Becker et al.’s and Snyder’s studies. The rigorous and well-documented research method included interviews, questionnaires, observations, and document analysis.

Findings. This study introduced the term ‘cue-consciousness’ as an extension of the notion of ‘the hidden curriculum’. Five out of the 30 students who responded in the study were ‘cue-seekers’ who actively sought out staff to ascertain their interests and what would be likely to be examined; 11 were cue-conscious and were perceptive to cues about assessment given by staff without going out of their way to seek these; 14 were cue-deaf and simply worked hard in order to succeed.

Limitations. This study was conducted in a very specific context involving a small group of elite students at an elite university, in a single discipline, studying under highly personalised conditions, and preparing for a particularly intensive form of high stakes exam.
**Generalisability.** The authors note that each department within the university is different in terms of assessment, and that universities differ from each other, yet claim that the issues they address are common. The sample is narrow and small, and the number of ‘cue-seekers’ (five) is dangerously so for anyone trying to generalise about cue-seeking behaviour beyond the sample studied. While the largest number of the sample were cue-deaf, this study is often cited to support the claim that students tend to be cue-seekers or cue-conscious.

Each of the above studies has made an important contribution to our understanding of student behaviour. *Making the grade* drew attention to the role of high achievement as a motivating factor. *The hidden curriculum* made the critical distinction between the formal curriculum as stated in course documents and the informal curriculum constituted by staff expectations which are not stated explicitly but are nevertheless ascertained by students who will look for them. *Up to the mark* introduced a useful set of terms to describe variation in the extent to which students will seek out these cues. At the same time, the numerous limitations of these studies noted above should make us cautious in applying their findings to contemporary students and higher education institutions:

- All of the studies were conducted at a time when university students were an elite group.
- Two of the studies (Snyder and Becker et al.) treat students as a homogeneous group, using methods that would not identify the sort of variation in student responses that we would now expect to find.
- Two of the studies were limited to one or two disciplines (Miller and Parlett – law; Snyder – science and engineering).
- These two studies were based on small student samples.
- The relevant aspect of the Edinburgh study focused on a very specific type of assessment – the final examination.

These studies have not been replicated, nor are they likely to be: they arose from a particular research paradigm, were more resource-intensive than is now normally possible, and their findings emerged only after several years’ work. Gibbs and Dunbar-Goddet (2007) note the absence of similar studies in their report on the effects of programme assessment on learning. The paper on ‘Conditions under which assessment supports students’ learning’ (Gibbs and Simpson 2004–2005), which provides the conceptual framework to the Gibbs and Dunbar-Goddet project, and the *Assessment Experience Questionnaire (AEQ)* developed from that framework represent an attempt, amongst other things, to operationalise the concepts of the hidden curriculum and cue-consciousness in wide-scale research across universities. Initial testing of the AEQ (Dunbar-Goddet et al. 2006; Gibbs 2006b) supports the proposition that assessment does influence students’ distribution of effort and their coverage of the syllabus, though the strength of this influence and whether it applies to some students more than to others is unclear. One research study, using a Chinese version of the AEQ and involving 108 education students in Hong Kong, produced contradictory findings, leaving it unclear whether students tended to believe that assessment allowed them to be selective in what they studied or required them to cover the entire syllabus (Joughin 2006). While the AEQ-based research and associated studies are still embryonic and not widely reported, they do promise to offer useful insights into how students’ patterns and foci of study develop in light of their assessment.
The influence of assessment on students’ approaches to learning

If assessment acts as a driving force of students’ study, it might be expected that different forms of assessment would affect study in different ways. Several studies concerned with this will be noted in this section. An earlier search for literature on this matter (Joughin 2008) identified a preponderance of studies within the ‘student approaches to learning’ framework and its predecessors. This may not be coincidental, since this framework shares with the above sociological studies a commitment to understanding the experience of learning and assessment from the student perspective, often providing a finer-grained analysis than that afforded by the sociological approaches. If assessment has a determinative influence on study, it might be expected that different kinds of assessment would lead to different approaches to learning – a proposition which studies in this tradition were methodologically well suited to explore. Nightingale and her colleagues are no doubt in good company when they conclude that ‘student learning research has repeatedly demonstrated the impact of assessment on students’ approaches to learning’ (Nightingale and O’Neil 1994, cited in Nightingale et al. 1996, 6), while numerous authors have cited with approval Elton and Laurillard’s (1979) aphorism that ‘the quickest way to change student learning is to change the assessment system’ (100).

The tenor of the student approaches to learning studies was set by the early work of Terry (1933) and Meyer (1934, 1935), followed by Silvey (1951), comparing how students studied for ‘objective’ tests and essay-style tests. Subsequent studies within the student approaches to learning tradition, including work by Thomas and Bain (1982, 1984), Tang (1994), and Scouller (1998), suggested that students’ approaches to learning are influenced by the form of assessment task they are undertaking or preparing for, with a growing recognition of the mediating function performed by students’ varying perceptions of the task. Two serious difficulties arise in relation to these and similar studies.

Firstly, many of these studies fail to provide important details of their context. Specifically, they typically fail to specify the cognitive demands of the assessment tasks under consideration. As a result, it is not possible to know if students are responding to a particular assessment format or to levels of cognitive demand associated with that format. For example, when students adopt a surface approach in preparing for a multiple-choice test, they may do so because they anticipate that the content being tested will be factual, or because the test format per se encourages a surface approach.

Secondly, to the extent that assessment may influence students’ approaches to learning, it may do so in one direction only. Struyven, Dochy, and Janssens’ (2005) review of 36 empirical studies is telling. While they begin with the proposition that ‘assessment is … one of the defining features of students’ approaches to learning’ (326), the studies they report lead them to conclude that ‘inappropriate assessment procedures encourage surface approaches, yet varying the assessment questions may not be enough to fully evoke deep approaches to learning’ (328). They note the research reported by Marton and Säljö (1997) almost 30 years earlier where attempts to induce deep approaches to learning had the opposite result and conclude that Marton and Säljö’s findings still stand:

It is obviously quite easy to induce a surface approach and enhance the tendency to take a reproductive attitude when learning from texts. However, when attempting to induce a deep approach the difficulties seem quite profound. (53)
This finding is strongly endorsed by Ramsden (1997) and Haggis (2003), leading to the conclusion that ‘the influence of assessment on approaches to learning may not be that more appropriate forms of assessment can induce a deep approach to learning, but rather that inappropriate forms of assessment can induce a surface approach’ (Joughin 2009, 22). Certainly the view that assessment, on its own, can be used to improve the quality of student learning is not supported by the evidence currently available.

A new research agenda

Notwithstanding the reservations noted so far in this paper, the conviction that assessment exerts a strong, even determinative, influence on our students’ learning continues to be widely maintained in higher education literature. Clearly, the empirical research outlined here provides only highly qualified support for this claim. Of course, this is not to say that the claim is not true, simply that the evidence cited in its support is far from unequivocal. The limitations of the research and the contrasts between the higher education contexts of the 1950s–1970s and the late 2000s in which we now teach and our students learn suggest that a renewed research effort is in order.

Perhaps the most fundamental question for such research is whether the very notion of ‘the hidden curriculum’ still applies. The higher education sectors at least in the UK, Australia, the USA, and Hong Kong have been strongly influenced by outcomes-based education and moves towards criteria-referenced assessment, with associated requirements to make curriculum objectives and assessment criteria explicit to students. It is possible that such developments have made the concept of the hidden curriculum less relevant. On the other hand, it should be noted that in Snyder’s study, his concern was not that expectations were not stated, but rather that students experienced a significant gap between formal statements of requirements and what teachers actually wanted from their students. Moreover, formal statements in course outlines and descriptions of assessment tasks, however detailed and carefully expressed, will only partly communicate clear expectations; they are likely to be perceived in different ways not only by different students but also by different teachers. These are empirical matters that are appropriate foci of research.

If the search for clarification of requirements continues, then the differential ways in which students approach this clarification through cue-seeking, cue-consciousness, or cue-deafness warrant further study. Most writers who cite Miller and Parlett’s work concentrate on ‘cue-seeking’ behaviour. A closer examination of the ‘cue-deaf’ student who is focused on learning rather than its assessment could be revealing: How do such students perceive the context of their study? How do they go about studying? Are they more appropriately focused on learning than strategically oriented students who are more attuned to meeting teacher expectations? Given the narrow sample of Miller and Parlett’s study, researchers could also be asking whether and how this kind of behaviour manifests itself in a range of disciplines, across year levels, and in different kinds of universities, and how it may be related to students various motivations for undertaking university courses.

Higgins, Hartley, and Skelton (2002) are rare in their questioning of the applicability of Becker et al.’s work to contemporary universities, noting that in the study of university life in America [sic], ‘assessment demands were ubiquitous, and student behaviour reflected the instrumental and pragmatic strategies they adopted to cope with the particular teaching and assessment strategies imposed on them’ and proceeding to ask, ‘But is this true for today’s students in the context of the UK?’ (59). Of
course, Becker and his colleagues were adamant that theirs was not a study of American colleges but of simply one possibly unique one, and thus they declined to generalise their findings beyond the very specific context of the University of Kansas. Consequently, the question posed by Higgins et al. needs to be asked in relation to all students, regardless of where they are studying. Equally unknown is whether contemporary students, with more complex social and economic lives than their earlier counterparts, are as dominated by the desire to achieve the highest grades possible as were their Kansas cousins in the 1950s and early 1960s.

While the overall impact of assessment on students study behaviour warrants a significant research effort, ongoing evaluation of the differential effects of various forms of assessment on learning may be equally important, particularly when the case for the so-called alternative assessment often depends on claims that they induce preferred approaches to studying. Moreover, in times of financial constraint and ever-increasing student numbers, moves to employ forms of assessment considered to be more efficient, such as examinations and multiple-choice tests, often proceed without consideration of how they may impact on learning. In such a context, ongoing research is needed to help determine whether specific assessment formats can: (1) merely support a pre-existing orientation towards deep approaches to learning, (2) induce deep approaches to learning in the absence of such a pre-existing orientation, and (3) induce surface approaches to learning where orientations towards deep approaches to learning are present. Studies utilising conceptual frameworks other than ‘student approaches to learning’ may provide important, different kinds of insights than those afforded by the kinds of comparative studies noted in this paper.

In pursuing questions such as these, the need for researchers to carefully and critically read the primary reports of past research now goes without saying.

Notes on contributor
Gordon Joughin is an associate professor in the Academic Development Unit at the University of Wollongong and visiting professor at Leeds Metropolitan University. His research and writing has focused on the relationship between assessment and student learning processes, with an emphasis on oral forms of assessment.

References


