# Accountability for Student Learning: Views from the Inside Out and the Outside In

**Victor M.H. Borden**

## Abstract

Because of an inherent and desirable diversity of opinion regarding what constitutes quality in postsecondary education, most quality assurance processes rely on institutions to define quality on their own terms in congruence with their missions. This approach allows for considerable autonomy and innovation but makes it difficult to convey to various constituents what specific institutions and the higher education sector as a whole contribute to society at large and to individuals and communities. This essay explores practical and conceptual issues related to increased demand for accountability for student learning outcomes through the U.S. lens of experience and offers a framework for a constructive approach to public accountability applicable to both the U.S. and South African contexts.

**Keywords:** accountability, Higher Education, educational quality, institutional effectiveness, organizational performance, educational policy.

## Introduction

Higher education institutions worldwide face increasing demands to demonstrate quality, educate a more diverse array of the world's population and contribute to the economic and social development of the state. These demands reveal the increasing importance of higher education to both individual and societal well-being. But they have also engendered a more diverse array of financial and political sponsors with varied interests and values in higher education. As a result, there has been a shift in some

oversight responsibilities and governance prerogatives from traditional self-regulation of the academy to an array of institutional management and external public and private sector interests.

Growing external demands on and interests in the work of higher education institutions overlay traditional intrinsic higher education values for preserving history, advancing culture, cultivating the life of the mind, and advancing scientific and technological frontiers. Academic administrators and administrative professionals find themselves navigating between two competing sets of interests and values: leveraging external demands to stimulate internal improvement while preserving autonomy and self-regulation important to academic professionals.

Despite significant differences among their histories and cultures, higher education institutions throughout the world experience these demands in markedly similar ways. Santos (2010) has characterized these competing demands in terms of three demands that arise from a specific value conflict: Production of high culture, critical thinking and exemplary scientific and humanistic knowledge for the training of the elites v. the production of average cultural standards and instrumental knowledge. All three demands are useful in the following ways. One demand is for training of the qualified labor force demanded by capitalist development. Another calls for hierarchization of specialized knowledge through restrictions of access and credentialing of competencies v. political demands for a democratized university and equal opportunity for children of the working class (or traditionally underserved populations). The third highlights the need for autonomy in the definition of unique values and objectives versus holding institutions to the same criteria of efficiency, productivity and social responsibility that private enterprises face. In South Africa, the most salient value conflict is between developing world-class, globally competitive research universities and higher education's role in redressing decades of oppression and exclusion among the African majority population.

The present analysis explores the tensions and paradoxes inherent in the current press for accountability and improvement in higher education institutions and systems primarily as experienced in the United States. The U.S. experience helps to bring into sharp relief the underlying tensions. On top of these the South African system experiences an even more vital and direct connection to the country's economic and social equity development.

Several organizing concepts are used to gain insight into the current context. In addition, a set of principles is offered to guide constructive approaches to navigating the turbulent social and political waters within which postsecondary institutions operate and seek support.

## Perspectives on Higher Education Quality

> Three factors distinguish top international universities from their competitors. The first: a high concentration of talented teachers, researchers and students …. The second factor … [is] their sizable budgets …. The third factor … is a combination of freedom, autonomy, and leadership …. [T]heir status is conferred by the outside world on the basis of international recognition[1].

> Learning—that is, the knowledge, skills and competencies a student gains by taking a college course or programme—really needs to be recognized as the primary measure of quality in higher education. Right now, that is simply not the case[2].

These two quotations about quality higher education contrast a traditional 'resources and reputation' view with a progressive 'student learning' perspective. Higher education institution leadership, professionals who guide academic programmes and support processes, and policy makers who seek to advance more effective accountability for higher education institutions will recognize in these perspectives some of the seemingly contradictory pressures that characterize public scrutiny of higher education. These perspectives apply especially in the United States, but increasingly throughout the world. Such perspectives raise the following questions. How important is it to improve an institution's position in popular rankings? Does doing so undermine difficult but critical efforts to promote attention to student learning outcomes? How can we improve degree completion rates

[1] Jamil Salmi, Tertiary Education Coordinator, World Bank (Salmi 2009).
[2] Jamie P. Merisotis, President, Lumina Foundation (Merisotis 2009).

while providing access to under-served segments of the population? Can we simultaneously develop improved capacities for research and innovation, and high quality learning experiences for more diverse learners? How can we keep college affordable when demands for advanced technologies, improved infrastructure, and higher service levels abound but public funding is increasingly scarce?

The 'resources and reputation' and 'student learning' perspectives appear to stand in stark contrast to each other. However, scratching below the surface reveals more nuanced issues related to reaching consensus on how colleges and universities should be accountable for the quality of their programmes, processes and outcomes. Consider the following statements:

> Virtually everyone who has thought carefully about the question of assessing quality in higher education agrees that 'value added' is the only valid approach. By value added we mean what is improved about students' capabilities or knowledge as a consequence of their education at a particular college or university[3].
>
> [Value added] implies that a student is doing just as well at an institution that graduates at the middle level of accomplishment (but with lots of improvement) as the student would do at an institution that graduates at the top level of accomplishment (but with less improvement)[4].
>
> [I]t should be obvious that quality is about content and intellectual innovation. If we are serious about having a high quality higher education system, then we have to start asking questions about content, avoiding the risk of suggesting that there is one standard way of measuring this[5].

[3] Douglas C. Bennett, President Earlham College (Bennett 2001).
[4] John V. Lombardi, President, Louisiana State University System (Lombardi 2009).
[5] Ferdinand von Prondzynski, President, Dublin City University (Von Prondzynski 2009).

> St. Ignatius Loyola … thought that the real test of higher education was what happened to the students—intellectually, socially, morally, and spiritually—under Jesuit tutelage. A university that measures its 'greatness' by application numbers and endowment rather than by the character of its graduates is a school with a decidedly secular notion of greatness[6].

> [T]he conversation about 'quality' has been centered on the wrong things. Institutional accreditation processes, despite their recent emphasis on assessing student learning and development, deal largely with resource and process measures. Government oversight as manifested in license requirements and programme review mechanisms, in turn, continues to emphasize regulation and procedural compliance. Third-party judgments of 'quality' such as media rankings continue to focus on such matters as student selectivity and faculty credentials. None of these gets at the heart of the matter: the investments that institutions make to foster proven instructional practices and the kinds of activities, experiences, and outcomes that their students receive as a result[7].

These statements provide a modest sense of the underlying diversity of opinion regarding what constitutes quality in postsecondary education. As a consequence, most institutional quality assurance processes rely on institutions to define quality on their own terms in congruence with their missions. Although this approach allows for considerable autonomy and innovation, it makes it difficult to convey to various constituents what specific institutions and the higher education sector as a whole contribute to society at large and to the individuals and communities they serve.

After nearly 50 years of increasingly intensive attention to this issue, we appear to be no closer to reaching consensus regarding how to characterize quality in higher education. This is not because we lack ideas as to what quality is, but because our diverse ideas, purposes, and interests

[6] George Weigel, Senior Fellow of the Ethics of Public Policy Center (Weigel 2005).
[7] National Survey of Student Engagement (2009).

make it difficult to establish standards and thereby to communicate effectively about higher education quality within our institutions as well as to a growing number of vested audiences.

## Motivations for Improving Quality

Why should members of the academy engage in efforts to evaluate and improve their programmes and processes? One answer we hear often is 'because we have to'. Beyond this reaction to the external pressure is a more fundamental reason: professional responsibility. Given the specialized knowledge and expertise required, academic staff have the rights and responsibilities of professionals to identify and hold themselves accountable to standards of competence and morality (Schön 1983). The strong and often confrontational calls for accountability from external sources emanates from mistrust and perceptions that members of professions are more self-serving than self-regulating. As a profession gains influence over the quality of life for individuals as well as for the state, the focal point of accountability shifts from members of the profession to the clients and their representatives.

This shift in focus for U.S. higher education is evident especially over the last 25 years. Although there were various 'clarion calls' for reform before this time, two seminal reports—*A Nation at Risk: The Imperative for Educational Reform* (National Commission on Excellence in Education, 1983) and *Involvement in Learning: Realizing the Potential of American Higher Education* (Study Group on the Conditions of Excellence in Higher Education, 1984)—touched off what has been a fairly continuous and increasingly acute focus on higher education accountability. Criticism regarding inadequate evidence of student learning, the focus of these reports, has been compounded by concerns about spiralling college costs and attendant concerns about affordability, opportunity, and gaps in participation and completion between traditionally served students and those from under-served populations. Over this same time period, a college credential has become recognized as increasingly important for both short- and long-term personal economic gain and higher education institutions recognized as increasingly critical to the technological, economic and social advancement of society through research, scholarship, and advanced professional education. Other societal trends, such as the increasing diversity of the

population and growing 'consumer sensibilities,' have added layers of complexity to the competing interests for accessibility and affordability on the one hand, and improved physical, technological and programmatic resources on the other.

The most recent decade, and especially the last five years was marked by a series of critical reports, regional and national commissions (e.g. the U.S. Secretary of Education Spellings' Commission on the Future of Higher Education), state and federal laws (e.g. the 2008 Higher Education Opportunity Act) and non-governmental organization initiatives to rein in higher education. In response to these pressures, academic associations and organizations have become further energized to both protect the academy and to advocate for reform from within. These associations and organizations seek to re-capture professional control and re-establish the trust necessary to work autonomously as self-regulated practitioners. Advocates for reform within the academy reason that the best way to support external improvement is to conduct systematic evaluation of academic programmes and student outcomes and use the results of that activity for programme improvement. For example, Ewell (2008a: 16) concludes that 'a genuine commitment to improve represents the best and most convincing evidence of accountability'.

Unfortunately, as Ewell also points out, conducting assessment for improvement purposes entails a very different approach than does conducting assessment for accountability purposes. Assessment for improvement entails a granular (bottom-up), faculty-driven, formative approach with multiple, triangulated measures (both quantitative and qualitative) of programme-specific activities and outcomes that are geared towards very context-specific actions. Conversely, assessment for accountability requires summative, policy-driven (top-down), standardized and comparative (typically quantitative) measures that are used for public communication across broad contexts. Information gleaned from assessment for improvement does not aggregate well for public communication and information gleaned from assessment for accountability does not disaggregate well to inform programme-level evaluation.

One could deduce from these differences that these two types of assessment should be pursued as independent activities. Unfortunately, decoupling assessment for improvement from assessment for accountability has undesirable consequences. Student learning assessment scholars and

practitioners have argued that assessment for improvement is most effective when it is embedded within the curriculum, and so has a direct connection to student learning. For example, Gray (2002: 61) notes, '[p]referred evaluation methods are those that are authentic, in that they fit naturally within the purpose and structure of a course … instead of only being used to grade students, the method also provides information to guide the improvement of both teaching and learning'. Curriculum embedded assessments are both practical and effective, since they do not require 'add-on' processes that have little or no value or meaning to students. Similarly, if assessment for accountability has little or no connection to assessment for improvement, then it is an 'add-on' process that adds significant overhead (and therefore contributes to escalating costs) without producing actionable results. Since improvement is the ultimate goal of assessment, it would be unproductive to decouple the two processes.

Paradoxically, assessment for accountability can undermine assessment for improvement, due to an inherent conflict in perspectives between the core audiences. Shulock (2005) describes this as an 'accountability culture gap'. Policy makers desire relatively simple, unambiguous information that provides clear evidence as to whether basic goals, such as programme completion and preparation for the workforce, are achieved. In contrast, Shulock (2005: 4) notes,

> The academic community finds bottom line approaches … threatening and inappropriate. [They] fear that such an approach can be punitive and can narrow society's concerns to those aspects of higher education that can be readily measured, at the expense of dearly held values. They fear legislative intrusion into matters of educational expertise …. They question how educational quality and equity can be quantified and assessed in a neat and tidy way and worry that quantitative measures create perverse incentives. They fear one-size-fits-all measures that ignore different missions, demographics, student bodies, resources, and factors outside their control. Most importantly, they resist legislative involvement in the measurement, or assessment, of student learning, which they believe to be a faculty responsibility.

Because of this culture gap, members of the academic community can be easily dissuaded from engaging with assessment for improvement when they see the kinds of measures of institutional effectiveness used within the accountability realm. For example, the strong emphasis in the U.S. on graduation rates among traditional students does not connect well with many members of the academy because of high levels of student mobility between institutions as well as connotations for lowering standards for performance. Although the mythical professor who tells students on the first day of class 'look at the students sitting to your right and left' to make the point that two of three students will fail the course is overstated, there is still a core value among members of the academy that students must earn their grades and that those who do not make the grade should not receive a degree. The senior academic administrators and professional staff that work to develop a culture of assessment within the institution can leverage core academic values to promote assessment for improvement. But their efforts are undermined by external emphasis on measures like graduation rates and their credibility can be challenged if they communicate these pressures either explicitly or through their resource allocation decisions.

Institutional leadership and faculty also find frustrating the ostensible conflict among the teaching/learning, research/scholarship, and outreach/service aspects of institutional mission. Various presses for improvement in each domain highlight the discrepancies and misalignments and downplay mutually reinforcing aspects, both of which are inevitable when pursuing multiple, complex missions. Although it is quite appropriate to question the appropriate balance in priorities and pursuits, the exclusive emphasis of any one aspect of mission and neglect of others contributes to the tensions that undermine credibility in pursuing improvement and accountability.

Ewell's (2008a) characterization of the two paradigms of assessment – improvement and accountability – helps us to understand the different pathways that need to be pursued in both domains and the accountability culture gap describes the chasm over which bridges need to be built to connect those pathways. To build those bridges and link accountability with improvement, we need to understand better and accommodate more effectively the relationship between the improvement and accountability domains. If not, we run the risk of engaging in costly, time-consuming

processes that do not effectively address either improvement or accountability objectives and only serve to widen and deepen the culture gap.

Several reports and studies in the U.S. have highlighted the important role of regional and specialized accreditation in stimulating the development of institutional assessment capacities (Aper, Cuver & Hinkle 1990; Harvey 2004; Wright 2003). Ewell (2009) describes how the shift of external stimuli over the last 20 years from state government to regional and specialized accrediting agencies provided a buffer between government agencies and institutions that promoted assessment practices more effectively to serve accountability and improvement simultaneously. However, he and others have noted that the relationship has not adequately satisfied public calls for accountability (Ewell 2008b; Ikenberry 2009; Neal 2008). In a recent U.S. survey of assessment practices, regional and specialized accreditation is described as the most important catalyst and motivator of assessment efforts, but current practices generally fall short of providing sufficient evidence for both accountability and improvement purposes (Kuh & Ikenberry 2009).

Because accountability is the primary stimulus for assessment but can also create significant barriers to effective practice, it is critically important for senior institutional leadership and the faculty and professional staff who coordinate and support assessment efforts to understand well the roots and manifestations of the relationship between the accountability and improvement functions. Campus leadership sets the tone for responding to demands for accountability as well as the need for assessment as a professional responsibility. As Ewell (2009: 15) notes, this can be approached as a compliance requirement or a common cause.

> Instead of seeing assessment as an aspect of higher education's responsibility to its funders—legitimate though this may be—both faculty and academic leaders need to see it as part of our accountability to ourselves. This is, after all, how we operate in the realm of research, and it is why mechanisms like independent peer review are so important to maintaining scholarly integrity. It needs to happen in teaching and learning as well.

## Paradoxical Tension and the Accountability/Improvement Dilemma

The accountability and improvement paradigms of assessment can be viewed as two sides of a coin: opposite yet inseparable. This idiomatic metaphor captures the existence dependence between the two domains of activity but does not represent how accountability pressures serve as both a catalyst for and a barrier to developing assessment for improvement capacities. The concept of 'paradoxical tension' provides further insight into the improvement/accountability relationship.

Paradoxical tension exists in a relationship when there are both conflicting and converging interests. Wagner (2009) proposes that paradoxical tension is fundamental throughout the biological and non-biological world and is also the basis for power and control over the social and physical world. The 'self/other' dualism is one such paradoxical tension. Individuals involved in a close relationship are often willing to abandon their self-interest in the interests of their partner, sometimes to the extreme that they would sacrifice their life for the safety of the other. Since partners are similarly disposed, neither wants the other to make such a sacrifice. The interests of parent and child are often paradoxically related in instances of protection (self-sacrifice for other) and punishment (inflicting short-term pain for longer-term gain). There are similar paradoxes between individuals and organizations, as between soldier and country: self-sacrifice serves one aspect of self-interest (devotion to country), while serving against another (self-preservation).

Within the realm of higher education, converging and conflicting interests of the self/other dualism are manifest within as well as across the improvement and accountability realms. Within the academy, they are manifest between instructor and student; programme chair and academic deans; academic deans and executive leadership. Within the accountability realm, they are manifest between executive leadership and trustee, between trustee and government agency heads, between agency heads and legislators; legislators and executives. The layers of complexity are particularly deep in the higher education realm, separating considerably the external oversight bodies from the core education and scholarship processes. This complexity helps to shield the processes at the lowest level but at the same time makes transparency extremely difficult as each layer adds a degree of opacity. There

are abundant instances of compatible and incompatible interests within the higher education sector and we don't always know which are which. For example, the relationship between scholarly productivity and undergraduate student learning can be seen as both mutually reinforcing activities and competing priorities.

Another dimension of the self/other dualism that relates closely to the accountability/improvement relationship is the 'actor/observer' difference in attributing causality (Jones & Nisbett 1972; Watson 1982), captured colloquially in the expression, 'I tripped but you fell'. The actor has an 'inside-out' perspective of his or her own behaviour and is likely to focus on external causes, such as a crack in the sidewalk. The observer takes an 'outside-in' view of the actor's behaviour, focusing on the actor's misstep as the cause. When attributing causality or blame in the case of negative results, the actor is more likely to focus on external forces and factors. The observer is more likely to attribute causality to the actor. This difference in perspective is not insurmountable, but it reflects tendencies toward causal attributions especially when actor or observer make only a cursory assessment of the situation without probing more deeply (i.e. taking the other's perspective into account or analyzing the situation more completely).

The actor/observer difference highlights the inside-out perspective of those engaged in assessment for improvement and the outside-in perspective in the accountability realm. Even when interests (aims and objectives) are common or at least compatible, there still can be (and usually is) tension related to differences in how each views responsibility for outcomes. From within the academy, problems are often seen as related to the materials with which and the environments within which the work occurs; that is, the attitude and behaviour of students and the availability of resources. The view from outside focuses on the behaviour of faculty and the quality of programmes and processes they enact.

Wagner (2009) also describes a 'matter/meaning' paradoxical tension that sheds further light on the nature of the accountability/ improvement relationship. This tension is closely related to the seemingly irreconcilable positivist and constructivist epistemologies. The mechanical (positivist) view of the world has had very practical uses in explaining how matter behaves in many circumstances. But there are serious flaws in this view discovered, for example, by quantum physicists. The 'meaning'

(constructivist) perspective, which posits that we create shared metaphors to help understand better what happens around us, is complementary in its ability to explain things that cannot be answered through a matter perspective but it too has limits and flaws. Wagner posits that such seemingly incompatible epistemologies are two sides of the same coin because each depends on the other for its existence (without matter there is no meaning and without meaning, no matter). However, we tend to see them as competing and incompatible views of the world and we use (and judge) them based on their ability to provide insights into what we come across in our lives.

The accountability perspective in higher education (and elsewhere) generally favours the mechanical, 'matter' point of view, presuming that there are basic 'facts' (graduation rates, levels of critical thinking, research productivity) that can be observed and compared across a broad array of contexts. Conversely, the improvement perspective generally takes a 'meaning' focus. Student progress takes on differing meaning depending on the structure of programmes and the concurrent obligations of the student population. Critical thinking is defined within disciplinary contexts or broader value contexts (i.e. local general education philosophies) and is assessed within the contexts from which it derives meaning. Similarly, research productivity depends upon disciplinary standards regarding modes of production and dissemination and any mechanistic view is likely to favour one discipline over another.

It is very difficult to resolve such paradoxical tensions because, as Wagner (2009: 150) notes:

> Neither party to a dispute sees that its entrenched position is just one side of a coin. It raises itself over other, whole over part, body over mind, matter over meaning, or vice versa and denies the other side legitimacy. Nothing may be wrong, except the belief that one of them is the truth.

Advancing our understanding in a world that is rife with paradoxical tensions, according to Wagner, requires that we choose one side for pursuing our inquiries but that we never forget that we have made a choice and that there is not an 'ultimate truth'. He also notes significant benefits to explicitly

acknowledging paradoxes, including the potential mitigation of hubris and the criticism and persecution of others who pursue their activities from different perspectives.

In the arenas within which higher education accountability and improvement play out, there are individuals who operate on the front lines of these paradoxes. Senior campus administrators and policy officials in government agencies; professional assessment and teaching and learning support staff and their counterparts within policy agencies; the leadership and professional staff within higher education associations and consortia and their counterparts at non-governmental agencies are among the front line workers on each side of the culture gap. These groups and individuals are responsible for promoting the pursuit of assessment for improvement and accountability simultaneously. As they promote the work necessary on each 'side,' they must work together to maximize the catalyzing and reinforcing components of the relationship while minimizing interference that arises from incompatible interests and differences in perspectives and approaches.

## Promoting Effective Assessment for Both Accountability and Improvement

Dealing effectively with the paradoxical tensions between the accountability and improvement realms requires that we understand clearly the differing viewpoints, accommodate the converging and conflicting interests and recognize the differing activities required to achieve core objectives. Although there is not likely an easy reconciliation, representatives can work together more productively acknowledging that each side has flaws and limits but both are worthwhile pursuits. A more productive engagement can occur by focusing on the integrity of work in both realms through guidelines and standards for effective, professional practice.

### *Principles for Accountability Frameworks that Promote Improvement*

Just as members of the academy should take professional responsibility for assessment as a vehicle for improvement and accountability, so to should

members of the policy domain take professional responsibility for the shape that public accountability takes and the impact it has on institutional and programme performance. Reporting on a forum sponsored by the American Enterprise Institute, Lederman (2009) concluded, 'if a major theme emerged from the assembled speakers, most of whom fall clearly into the pro-accountability camp, it was that as policy makers turn up the pressure on colleges to perform, they should do so in ways that reinforce the behaviours they want to see – and avoid the kinds of perverse incentives that are so evident in many policies today'. If accountability requirements are the primary catalyst for assessment efforts within institutions then they also shape and sometimes distort the tenor and nature of those efforts. Four principles are offered for crafting accountability frameworks that induce more authentic institutional efforts to identify, assess and improve quality.

> Defining Quality. Effective accountability starts with the articulation of specific quality objectives that accommodate the diverse core objectives of higher education (e.g. access and affordability, learning, research and scholarship, community engagement; technology transfer, cultural enhancement and so on).

The quality of student learning was the central focus of the seminal reports that catalyzed the current focus on higher education accountability, as well as many of the reform efforts from within the academy since that time. The traditional ''reputation and resource'' view has been criticized as inappropriate but it has not abated. Some have suggested that media-based rankings promote this outmoded view beyond its usefulness while others suggest that the popularity of these rankings indicate that there is still merit in this view of quality. While this debate continues, advocates of other aspects of institutional quality, such as equity in participation and performance, student character development and the civic engagement of institutions in their communities, seek recognition for their causes. Student learning within undergraduate-level programmes is a nearly-universal and undeniably important enterprise across the higher education landscape that deserves acute attention. Because of its pervasiveness and complexity, it is important to recognize that student learning outcomes cannot be reduced to a few quantifiable measures, lest we reduce incentive for faculty to engage

authentically in assessment processes. It is essential that we accommodate both the diverse range of student learning objectives evident across the US higher education landscape, as well as other mission-critical purposes that differentiate and distinguish postsecondary institutions.

> Accommodating Diverse Learners. Accountability frameworks should explicitly recognize differences according to the population spectrum that is served by institutions and programmes, and should do so in a way that does not suggest that there is greater value for serving one segment of the population over another.

Quality across various higher education functions (e.g. learning, scholarship, engagement and outreach) is further shaped by the spectrum of the population that participates in or benefits from that function. Using common measures and standards to compare institutions that serve markedly different student populations (e.g. a highly selective, residential liberal arts college compared to an open-access, community college with predominantly part-time students, or comprehensive public university serving a heterogeneous mix of students) results in lowered expectations for some types of institutions and unreasonable demands for others. If similar measures are used but 'acceptable standards' are allowed to vary, an inherent message is conveyed that one type of mission is inherently superior to the other. The diversity of the U.S. higher education landscape is often cited as one of its key strengths. Homogenous approaches to quality assessment and accountability work against that strength and create perverse incentives that undermine important societal goals. For example, there is a growing body of evidence that the focus on graduation rates and attendant concerns with student selectivity (the most expeditious way to increase graduation rates) has incentivised higher education institutions as well as state systems to direct more discretionary financial aid dollars to recruit better students rather than meet financial need. This, in turn, has reduced the proportions of students from under-served and low-income families that attend four-year institutions and that complete a college degree (Bracey 2005; Dynarski 2002; Heller 2006; Gold & Albert 2006; Mortenson 2009; Ness & Noland 2007).

Programmes and institutions should be held accountable for their particular purposes and on the basis of whom they serve. Those who view

accountability from a system-level perspective should recognize explicitly how institutional goals differentially contribute to broader societal goals by virtue of the different individuals and objectives the institutions serve. Promulgating common measures or metrics, or at least comparing performance on common measures, does not generally serve this purpose.

> Connecting Performance with Outcomes. Accountability frameworks should facilitate making connections between performance (programmes, processes, and structures), transformations (student learning and development, research/scholarship and professional practice outcomes) and impacts (how those outcomes impact the quality of life for individuals, communities, and society at large).

Once the basis for quality (what for whom) is better understood and accommodated, we can assess for both improvement and accountability purposes, how various programmes, structures, organizations and systems contribute to the production of quality education, research and service. To do so it is helpful to distinguish among three inter-related aspects for our measures and inquires:

a) Performance. How well do the processes, programmes, and structures of an institution perform in relation to professional standards? This is a traditional focus of the accreditation realm, examining the resources available for performance and the policies and processes enacted.

b) Transformation. What happens to people and bodies of knowledge to which these resources and processes are applied within our higher education organizations and systems? Student learning outcomes are one component of this focus as are other transformations sought through scholarly research and public and professional service. Boyer's (1997) model of scholarship provides an integrated description of the range of transformations that individuals and institutions in the higher education sector pursue: discovery; integration, application, and teaching/learning.

c) Impacts. The near-term transformations pursued are, in a sense, intermediary outcomes for longer-range objectives. Better educated students lead more productive lives, economically and civically. Advances in science and technology result in products and professional practices that improve the quality of life. The preservation and advancement of arts and culture contributes in different but equally important ways to the quality of life. The quality of higher education programmes and processes relates closely to how short-term outcomes link to these longer-term objectives as such linkages are not always evident and not always positive.

Although these three manifestations of quality—performance, transformation and impacts—are inextricably intertwined, it is useful to keep in mind the distinctions in formulating assessments for improvement and accountability measures. Efforts to improve higher education require that, within the academy, we understand better how our structures, programmes and processes perform to produce desired transformations that result in positive impacts. Accountability, as an external catalyst for improvement will work best if we reduce the perverse incentives that arise from measures that do not connect appropriately among the aspects of performance, transformation and impact sought by the diverse array of postsecondary organizations and systems that encompass the national and international higher education landscape.

Validity for Purpose. Accountability measures should be assessed for validity related specifically to their intended use, that is, as indicators of program or institutional effectiveness.

Wagner's (2009) thesis on paradoxical life further informs this discussion as he suggests that, in the biological, social and even hard science realms, we are at a point in history where we need to pay more attention to meaning than to matter. Accountability frameworks tend to focus on measures without sufficient attention to meaning. Measures like graduation rates and expenditures per full-time equivalent student are not consistent and reliable across different contexts but are accepted as the coin of the realm. As we now seek to discover more meaningful measures of student learning and

other important transformations and their impacts, it is critical that we tend to the meaning behind measures.

Reliability and validity are the quintessential criteria for effective measurement. Reliability refers to the mechanical (matter) aspects of measurement, that is, the consistency of a measure or assessment within itself and across differing conditions (Borden & Young 2008). Validity refers to the relationship between the measure and meaning. The concept of validity has evolved throughout the 20th century with more recent focus, as shaped especially by Cronbach (1988) and Messick (1989), on 'how strongly theory and evidence support the interpretations and decisions that are based on a measure' (Borden & Young 2008: 21). By this view, a measure does not have validity in and of itself, but rather only in terms of the assertions made using the measure. For example, one cannot say that a specific measure used as a graduation rate (most commonly, in the U.S. the proportion of first-time, full-time degree seeking college students who complete a degree within six years) is or is not valid. Rather, one can argue that it is or is not a valid measure of institutional effectiveness (or student success) within varying contexts (e.g. at a highly selective, national liberal arts college compared to an open access regional commuter university).

Borden and Young (2008) discuss the current poor state of validity assessment in the realm of higher education accountability measures. They use as an illustrative case the lack of evidence for validity of the value-added measure of critical thinking and reasoning that is included as a pilot project within the Voluntary System of Accountability (VSA http://www.voluntarysystem.org/). According to Borden and Young, there is an important difference between the validity of a test or measure as an indicator of student ability and the validity of the same measure as an indicator of institutional effectiveness (i.e., that the institution positively contributed to the student's ability). In addition, because there is far more variation within than there is across institutions with regard to student performance, the use of an institutional average or value-added score as a measure of institutional effectiveness is inappropriate unless there is direct evidence that institutional programmes and processes influence the transformation of student knowledge, skills and abilities and do so more than other factors that account for score differences across institutions (e.g. differences in student body characteristics, programme mix, consistency of measurement methods, etc.).

The recent validity report (Shulenberger & Keller 2009) for the three options for standardized measures that can be used within the VSA includes some circumstantial evidence related to use of the measure at the institutional level, specifically that seniors score consistently higher than freshmen, on average, across the exams. However, as Shulenberger and Keller point out, the study did not directly evaluate the value-added score that will be used in the VSA pilot project.

Borden and Young (2008) describe a set of standards for validating accountability measures that begin with a description of the kinds of inferences and claims that are intended to be made with the measure, for example, that students at institutions with higher average (or value-added) scores on a standardized exam learned more than did students at institutions with lower scores. The standards then focus on the conceptual basis for these claims, for example, how is learning construed and expressed through the exam scores? The third standard focuses on the basis of evidence that is sufficient for backing the claims. The last two standards relate to the methods used to collect evidence, that is, whether implementation is appropriate from a technical perspective and whether it is fair, given the contexts within which measurement is taken and potential differences in applicability of the measure to differing groups and populations.

There is generally little if any attempt to ensure that accountability measures support the claims that are intended by their use. This is not surprising, given the processes that are used to develop accountability measures. Often (at best), significant thought, negotiation and technical review go into designing measures. However, there is generally little done to empirically assess the validity of the measures in relation to the inferences and claims that are made using them. Establishing validity of accountability measures is an 'inside-out' perspective that needs to be applied more to the accountability realm. Those who promulgate accountability need to take professional responsibility (and be held accountable by members of the academy) for establishing validity of required measures and methods.

The state of validity assessment within the higher education realm (and education more generally) contrasts starkly to the more stringent requirements for validity imposed within the scientific research and health domains. Because there is a close interdependence between accountability demands and the shape and tenor of institutional efforts to assess and

improve quality, it is critical that accountability frameworks employ the same level of professional competence that is expected of institutional efforts.

## Conclusion

It is not possible to completely reconcile the paradoxical tensions between externally-oriented accountability demands and efforts within higher education institutions to meet challenges of educating a broader spectrum of the population and contribute to economic and social developments while containing costs. However, it is possible to advance efforts in both spheres if we recognize the inherent paradoxical tensions and accord the individuals pursuing these efforts the rights and responsibilities for doing so. Members of the academy should accept the imposition of accountability standards, recognizing the increasing importance of higher education to a broader range of vested interests. At the same time, the academic community and others should hold those invoking accountability (government agencies, non-governmental agencies and the media) to professional standards so as to promote positive (and not perverse) incentives for pursuing core objectives. Those seeking more accountability, in turn, should recognize that a 'one-size-fits-all' approach to accountability does not accommodate well the increasingly diverse landscape of higher education providers and the diversity of the populations served.

Although notable attention has been paid to higher education accountability as a general issue (e.g. Leveille 2006; Mortimer 1972; National Commission on Accountability 2005; Vogel 2006), the quality of assessment and measurement in the accountability realm has not received nearly as much attention, especially from the perspective of professional responsibility. The individuals and entities involved in developing accountability frameworks and measures are often embedded within political and policy realms where final decisions are influenced by political agendas and expediencies that do not accommodate complexity and nuance. Similarly, the rankings and ratings that appear in public media are motivated first and foremost by sales, which, in turn, depend on the perceptions of the consuming public regarding what comprises quality and what makes a compelling story. The rapid expansion of available instruments for assessing higher education programmes and services, especially from for-profit

providers, indicates further that assessment is being shaped as a commodity that can be bought and sold.

Addressing the complex challenges required to expand the benefits of higher education to more individuals and more communities requires a careful balancing between pressures to demonstrate effective practices and the latitude and autonomy that promotes innovation and creativity. Ill-conceived and haphazard approaches have not proven effective in either the internal improvement or external accountability domains. Progress is more likely to be achieved through a constructive engagement that leverages the intrinsic motivations of professionals taking responsibility for their own work while acknowledging the valid interests among external constituents for cogent evidence of the value obtained for investment and support.

In the South African context, the nuances of these 'accountability problems' may seem to be relatively trivial matters in relation to the deeper and more intractable societal issues that institutions of higher education are expected to address. However, these accountability demands underlie increasing expectations for demonstrable performance improvements according to the measures used to fund and advance institutional development. South African institutions that can successfully balance these competing societal pressures and demonstrate the value added to student and societal development have a greater chance of being recognized and supported within the country as well as globally, given the relatively small number of institutions within the country relative to its size and visibility. Similarly, the South African higher education system can promote quality improvements more effectively if policy makers and others who set the overall accountability agenda take the professional and constructive approach to defining accountability requirements proposed in this essay.

## References

Aper, JP, SM Cuver & DE Hinkle 1990. Coming to Terms with the Accountability versus Improvement Debate in Assessment. *Higher Education* 20: 417-383.

Bennett, D 2001. Assessing Quality in Higher Education. *Liberal Education* 87,2:406.

Borden, VMH & JW Young 2008. Measurement Validity and Accountability for Student Learning. In Borden, VMH & GR Pike (eds): *Assessing and Accounting for Student Learning: Beyond the Spellings Commission*. San Francisco: Jossey-Bass. (New Directions for Institutional Research, No. S1.)

Boyer, EL 1997. *Scholarship Reconsidered: Priorities of the Professoriate*. San Francisco: Jossey-Bass.

Bracey GW 2005. The Bachelor's Degree: A Hereditary Privilege? *The Phi Delta Kappan* 87,3: 253-254.

Cronbach, L 1998. Five Perspectives on Validity Argument. In Wainer, H & H Braun (eds): *Test Validity*. Mahwah, NJ: Erlbaum.

Dynarski, S 2002. *The Consequences of Merit Aid. NBER Working Paper Series.* Cambridge MA: National Bureau of Economic Research. Available at: http://papers.nber.org/papers/w9400.pdf. (Accessed January 6, 2012.)

Ewell, PT 2008a. Assessment and Accountability in America Today: Background and Context. In Borden, VMH & GR Pike (eds): *Assessing and Accounting for Student Learning: Beyond the Spellings Commission*. San Francisco: Jossey-Bass. (New Directions for Institutional Research, No. S1.)

Ewell, PT 2008b. *U.S. Accreditation and the Future of Quality Assurance*. Washington, DC.: Council for Higher Education Accreditation.

Ewell, P 2009. *Assessment, Accountability, and Improvement: Revisiting the Tension.* Urbana, IL: University of Illinois and Indiana University. (National Institute of Learning Outcomes Assessment, Occasional Paper No. 1.) Available at: http://learningoutcomesassessment.org/documents/PeterEwell_005.pdf. (Accessed January 6, 2012).

Gray, PJ 2002. The Roots of Assessment: Tensions, Solutions and Research Directions. In Banta, TW & Associates (eds): *Building a Scholarship of Assessment*. San Francisco: Jossey-Bass.

Gold, L & L Albert 2006. Graduation Rates as a Measure of College Accountability. *American Academic* 2,1:89-106.

Harvey, L 2004. The Power of Accreditation: Views of Academics. *Journal of Higher Education Policy and Management* 26,2: 207-223.

Heller, DE 2006. *Merit Aid and College Access. Symposium on the Consequences of Merit-Based Student Aid, Wisconsin Center for the*

*Advancement of Postsecondary Education.* Madison, WI: University of Wisconsin. Available at: http://inpathways.net/meritaid_college access.pdf. (Accessed January 6, 2012.)

Ikenberry, SO 2009. *Where do we Take Accreditation? White Paper Prepared for the 2009 Council for Higher Education Accreditation Annual Conference.* Washington, DC. Available at: http://www.chea.org/pdf/2009_AC_Where_Do_We_Take_Accreditation_Ikenberry.pdf. (Accessed January 6, 2012.)

Jones, EE & RE Nisbett 1972. The Actor and the Observer: Divergent Perceptions of the Causes of the Behavior. In EE Jones, DE Kanouse, HH Kelley, RE Nisbett, S Valins & B Weiner (eds): *Attribution: Perceiving the Causes of Behavior*. Morristown, NJ: General Learning Press.

Kuh, G & S Ikenberry 2009. *More than you Think, Less than we Need: Learning Outcomes Assessment in American Higher Education.* Champaign, IL: National Institute for Learning Outcomes Assessment (NILOA), University of Illinois at Urbana Champaign. Available at: http://www.learningoutcomeassessment.org/documents/niloafullreportfinal2.pdf. (Accessed January 6, 2012.)

Lederman, D 2009. Defining Accountability. *Inside Higher Education* November 18. Available at: http://www.insidehighered.com/news/2009/11/18/aei. (Accessed January 6, 2012.)

Leveille, DE 2006. Accountability in Higher Education: A Public Agenda for Trust and Cultural Change. Berkeley, CA: Center for Studies in Higher Education, University of California. Available at: http://www.cpec.ca.gov/CompleteReports/ExternalDocuments/Leveille_Accountability.20.06.pdf. (Accessed January 6, 2012.)

Lombardi, JV 2006. Virtues and Vices of 'Value Added'. *Inside Higher Education* August 10, 2006. Available at: http://www.insidehighered.com/views/2006/08/10/lombardi. (Accessed January 6, 2012.)

Merisotis, JP 2009. It's the Learning, Stupid. The Howard R. Bowen Lecture, Claremont Graduate University, Claremont, CA, October 14, 2009. Available at: http://www.luminafoundation.org/about_us/president/speeches/2009-10-14.html. (Accessed January 6, 2012.)

Messick, S 1989. Validity. In Linn, R (ed): *Educational Measurement.* 3rd Edition. New York: Macmillan.

Mortenson, TG 2009. *Family Income and Educational Attainment, 1970 to 2008*. Oskaloosa, IA: Postsecondary Education Opportunity No. 209.

Mortimer, KP 1972. *Accountability in Higher Education. American Association for Higher Education and the ERIC Clearinghouse on Higher Education.* Washington, DC: ED058465.

National Commission on Accountability in Higher Education 2005. *Accountability for Better Results: A National Imperative for Higher Education.* Boulder, CO: State Higher Education Executive Officers. Available at: http://www.sheeo.org/Account/accountability.pdf. (Accessed January 6, 2012.)

National Commission on Excellence in Education 1983. *A Nation at Risk: The Imperative for Educational Reform.* Washington, DC: United States Department of Education. Available at: http://www.ed.gov/pubs/NatAtRisk/index.html. (Accessed January 6, 2012.)

National Survey of Student Engagement 2009. Our Origins and Potential. Available at: http://nsse.iub.edu/html/origins.cfm. (Accessed January 6, 2012.)

Neal, AD 2008. Seeking Higher Education Accountability: Ending Federal Accreditation. *Change* 40,5: 24-29.

Ness, EC & BE Noland 2007. Targeted Merit Aid: Implications of the Tennessee Education Lottery Scholarship Program. *NASFAA Journal of Student Financial Aid* 37,1:7-17.

Salmi, J 2009. What Makes a University Great? *Forbes Magazine* August 8, 2009. Available at: http://www.forbes.com/2009/08/10/world-class-best-university-ranking-world-bank-opinions-colleges-salmi.html. (Accessed January 6, 2012.)

Santos, Boaventura de Sousa 2010. The University in the Twenty-first Century: Towards a Democratic and Emancipatory University Reform, Eurozine, October. Available at: http://www.eurozine.com/pdf/2010-07-01-santos-en.pdf. (Accessed January 6, 2012.)

Schön, DA 1983. *The Reflective Practitioner: How Professionals Think in Action.* London: Temple Smith.

Shulenberger, D & K Keller 2009. *Interpretation of Findings of the Test Validity Study Conducted for the Voluntary System of Accountability.* Washington, DC: Voluntary System of Accountability. Available at:

http://www.voluntarysystem.org/docs/reports/VSAabstract_TVS.pdf. (Accessed January 6, 2012.)

Shulock, N 2003. *A Fundamentally New Approach to Accountability: Putting State Policy Issues First. Institute for Higher Education Policy and Leadership.* Sacramento: California State University. Available at: http://www.csus.edu/ihe/PDFs/ASHE%20paper%202003%20FINAL.pdf (Accessed January 6, 2012.)

Study Group on the Conditions of Excellence in American Higher Education 1984. *Involvement in Learning: Realizing the Potential of American Higher Education.* Washington, DC: National Institute of Education. ED246833. Available at: http://www.eric.ed.gov/PDFS/ED246833.pdf. (Accessed January 6, 2012.)

Vogel, M (ed) 2006. Accountability: To Whom and for What? *American Academic* 2,1: 1 - 135.

von Prondzynski, F 2009. A University Blog: Diary of a University President. Available at: http://universitydiary.wordpress.com/2008/10/09/what-is-quality/. (Accessed January 6, 2012.)

Wagner, A 2009. *Paradoxical Life: Meaning, Matter, and the Power of Human Choice.* New Haven, CT: Yale University Press.

Watson, D 1982. The Actor and the Observer: How are their Perceptions of Causality Divergent? *Psychological Bulletin* 92: 682-700.

Weigel, G 2004. What Makes a University 'Great'?' *The Catholic Difference* (July 6, 2004). Reprinted with permission by the Catholic Education Resource Center. Available at: http://www.catholiceducation.org/articles/education/ed0266.html. (Accessed January 6, 2012.)

Wright, BD 2003. Accreditation and the Scholarship of Assessment. In Banta, TW & Associates (eds): *Building a Scholarship of Assessment*. San Francisco: Jossey-Bass.

Victor M.H. Borden
Professor of Educational Leadership and Policy Studies
Indiana University, USA
vborden@indiana.edu