Ways to Play the Outcomes Game

Russ A. Schultz
Lamar University

When I was a youngster, there was a joke that there were only two requirements in life, Death and Taxes. If you are in the field of education, there are three others, accountability, assessment and outcomes, words that are often used interchangeably. While we have all been overwhelmed and over consumed with assessment and evaluations, we should not lose sight that there are models that are sound and that the process, when done correctly, can be beneficial. In spite of what you have come to believe, all assessment activities are not problematic. Some institutional or academic association models set out to help us define what our students should know and be able to do, rather than simply reinforcing what we teach. One of the most prominent aspects of the “No Child Left Behind” Act recognizes that we must assess content or outcomes rather than trying to measure methodology. That is a positive change. Over the years this type of protocol has been called many things, competency-based instruction, outcome based instruction or under the rubric of the many assessment models. It is interesting that the arts have always embraced this type of review and been a strong advocate for content assessment regardless of the nomenclature or popular jargon of the time.

When you assign a particular artistic project to a student, there is an expectation of what that student will learn by the end of the process. If conceived and applied properly, this process will also provide us with the feedback to determine if the educational content has been consistent with the goals and if the student has acquired the skills that were originally delineated. Specifically, if the student succeeds, that gives the instructor one type of feedback or assessment of the student’s skill. However, if the student does not succeed, the instructor can question the student’s ability or, on the other hand, whether the program has provided sufficient data and/or skill development that would be necessary to achieve the desired results. In short, in the best assessments we are looking to validate that what is actually learned at the end of the activity was what we set out to accomplish. Problems occur in the process when we try to quantify these outcomes and then develop comparisons that truly defy logic.

With regard to the principles supporting assessment, that is, evaluating student outcomes and competencies, I am pleased to tell you that there is both good news and bad news. The good news is that you have been doing it for years. The bad news is that the responsibilities for establishing consistent and relevant programmatic outcomes, which were previously controlled by you, the practitioners, have now been taken out of your hands. Historically, those of us directly involved with instruction in the discipline have established the outcome or performance standards. We were charged with and assumed responsibility for determining what a student must know and be able to do. Then through a variety of evaluative tools, we determined if the student had acquired the skills that were generally recognized to be necessary and important. Over a period of time and encompassing the review of a large body of student evaluations, one could determine if there was consistency between the curriculum and student outcomes. For example, an annual review of the body of your students’ work is a reasonable assessment tool. Through this process one could ascertain the individual’s skill level and also the success of the defined curriculum. If the curriculum dictates the acquisition of a particular artistic skill or competency in a medium and all but one student attained this mastery, we can assume the problem to be with the student.
However, if all of the students lacked mastery of this skill, there would appear to be either a curricular or instructional problem related to the attainment of this skill.

In theory, that is how the process should work. However, in many cases there is a breakdown because of several factors. This session is to provide assistance in responding to the complexities of assessment that are being thrust upon us. While that is a focus of my presentation, it would be helpful to identify and understand the pressures being applied by either political or some regional accrediting bodies.

Over 40 states have begun tying some type of assessment to institutional funding; or what some politicians call “performance based funding.” In the current political climate, there is a perceived need to have greater accountability in higher education. This sound bite mentality, it is believed, will develop a learner that is more competitive on a global scale. However, the call for greater accountability is often connected to an understanding of higher education that is both fragmented and based on data that is, at best, reliably questionable. To be precise, an understanding of higher education based on limited individual experiences or anecdotal information and data that may have a basis in fact but is for the most part incomplete. It then becomes the goal of the accountability exercise to affect change but, based on fragmented information and questionable data, there is no clear direction and the probability of successful results is rather slim. In addition, often, in the body politic, education is one of the few places where legislators can exercise direct budgetary control, as other areas of the state budget appear to have less wiggle room. As a result, K-12 and higher education are used as political footballs to be kicked back and forth. Political control of funding in education is more successful and better accepted as the population has more contact with and understanding of these areas of the state budget. Therefore, politically, on the surface, the population believes these assessment models are responsible actions. However, these actions rarely, if ever, produce a better student or a higher level of accomplishment.

With appropriations, some legislators consider that they have made considerable investments in higher education and, therefore, have a responsibility to exercise control. However, the hand that gives can also take away and, using budgetary allocations as both the carrot and the stick, they can regulate assessment procedures that empowers greater jurisdiction over the process and its outcomes. As a means of budgetary control, this process generally employs two different approaches. The first is the assessment of only non-specific or non-qualitative items, such as classroom size or credit hour production. These can be easily quantified but are difficult to directly connect to quality. The second, which is less common, is the attempt to quantify qualitative outcomes. It was reported that Einstein had the following quote on his wall: “Not everything that counts can be counted, and not everything that can be counted counts.”

It is easy to understand that whoever is empowered to establish the criteria, develop the assessment tools, and delineate the comparative data will also control the outcome and consequently have their desired viewpoint validated. As a result, in this controlled culture it can be demonstrated that higher education is either working or not working in a manner that is consistent with the prevailing political view.

In order to legitimize the process to the public, it is commonplace to hold a small portion of funding back and provide the means by which to return it to the institution, so that the process seems more like a reward than a punishment. While the ultimate goal of change may be accomplished, it was not the people who best understand the academic environment who controlled the alterations.

The assessment and accountability process has developed as a parallel to a business model. However, educational quality cannot be measured in a manner that parallels the marketing of a product or service. Ultimately, the manufacturer has oversight of the product or service with checks and balances from the free market. If the product or service is good and needed, in theory it should sell. However, in this assessment model, the one area that has historically been the bastion of faculty jurisdiction, the curriculum, has been taken over by others. We are held responsible for the results without full jurisdiction over the content or process. While the resultant funding controlled by the assessment process
is usually only 1 or 2 percent, it is significant enough to provide real financial incentive to “teach towards the test,” or redefine the curriculum to better reflect the desired results. As a consequence, those controlling and developing the tests or assessment instruments take control of the curriculum; this is by far the most serious change caused by this process. Therefore, in an attempt to retain oversight of the curriculum, we have become expert in manipulating the data rather than doing what is actually conceived as the structural variants to achieve the desired outcome. Often if we do not like the results, we revise the input and manipulate the data to affect the outcomes, rather than making the necessary substantive changes that the assessment was trying to dictate. The assessment model, good or bad, has forced institutions to make changes that ultimately affect the curriculum. In other words, we have learned to dance around the model.

Difficulties have arisen with the desire to quantify quality in a manner that makes quantification difficult, if not impossible. There are things, such as artistic quality, that can’t be quantified. The basis of this trend is a belief that if we do it, we should be able to transcribe it into some digital quotient and with that number, be able to compare it to others. It appears that many in the political arena desire to develop the means by which they can make judgments based on some numerical consistency that they believe can simplistically tell the story. In other words, give me a number that tells me what I want to know and from that number I can make decisions. Clearly then, the department that scores a 2 is twice as good as the department that only scores a 1. Add to the confusion the lack of understanding of higher education in general and particularly the arts and we can understand why legislatures or accrediting bodies try to evaluate and assess characteristics that are not related to student learning but rather some characteristic that is subordinate to it.

As an example, while I was in the State of Washington, the legislature became concerned with student learning in higher education. However, they chose to assess this through three different criteria unrelated to the actual learning process. The first was to determine the credits a student would take in order to complete their degree. One of the criteria that the legislature evaluated was something that was called, Graduation Efficiency. They even devised a Graduation Efficiency Index number that calculated the number of credits required for a degree, divided by the number of credits that the student actually took to complete the course of study. Of course the legislature wanted the resultant number to be 1. In its wisdom, the legislature required that the university attain an average Graduation Efficiency Index not lower than .95. Any institution with a number lower than .95 meant that the students were wasting taxpayer’s money by taking all of those extra classes and receiving all of that additional knowledge, G-d forbid. There was a feeling communicated by this action that all of the students should enter the institution academically prepared, have decided on a major and therefore require no additional preparatory, supplemental or elective course work. This was a much more efficient view of student learning than was exercised in practice. These “efficient” students could and should graduate in four years or less and therefore reduce the waste of resources. Lost in this quotient were the desires of educators and the mission of the institutions to cultivate greater learning. There was no expectation that students might change their mind or possibly change their major.

The legislature saw the credits produced by students above what is delineated by their degree plan as something that should be discouraged. In other words, with this same outcome there were two very different interpretations. On the side of the faculty, they were pleased that students pursued greater learning, while the legislature equated learning with credits, not knowledge, and wanted to reduce the number of what they felt were extra credits.

Now most of us would agree that the desire to get the students to complete their course of study more quickly is a very positive goal. Some students are very dedicated in their attendance year after year. It would be quite positive and appropriate to move them along and out of the institution. However, this described process, with only the quantitative outcome, did not allow students to develop their educational and curricular goals appropriately, resulting in practices that were contrary to accepted models. Because
the legislative interpretation of the desired outcome in this example was so catastrophically different than
the recognized institutional and educational mission, it was difficult for reasonable people to understand
why the former view even existed. I will tell you that in this case we did not, as the legislature desired,
reduce the number of courses taken by the students for graduation. Instead, we modified the delineated
curriculum to ensure that all classes, especially those normally taken as electives, were now in the course
of study, ever enlarging the minimum requirements to attain the degree. By increasing the delineated size
of the program, we were able to continue to offer and allow students to enroll in needed courses and keep
the Graduation Efficiency Index in check. In other words, we manipulated the data.

It would be helpful for us to see that even in the most bizarre assessments, there may be some merit and
we should work to accept and pursue those areas. In a case like this, where art students often take large
numbers of classes and studios, exceed the minimum graduation requirements by many credits and still
produce a low number of student credits for the faculty, it is sometimes helpful to demonstrate
improvement over time. That is, over a period of time, if you cannot successfully meet the target number,
at least be able to show progress towards it.

In addition to the Graduation Efficiency there was also an Undergraduate Student Retention Index which
had a target of .95 and Five year Graduation Rate Index which had a target of .65. Approximately 2% of
institutional funding was held back with the possibility of regaining these funds based on the institutional
successes with these factors. None of these factors actually addressed quality, but were subordinate to
instructional activities. It is important to continue to remind ourselves that we are in the business of
developing quality students, even through the fog of these simplistic governmental controls.

In assessment there is a certain euphoria attained by those in control suggesting that one size fits all. That
is, if we can quantify something, we can compare it. As long as we can generate a number there is reason
to think that what works in chemistry will also reflect effectively the work in art, dance, theatre and
music. In the academy where we tend to exalt diversity, we are forced to accept that through external
intervention, the assessment of one discipline can, without question, be completed and compared in the
same manner as assessments in another discipline. This is done regardless of the make-up of the
discipline or the resultant overall success of the assessment plan. The purpose of these less than accurate
evaluations is that the person/organization mandating the assessment can in some way make a
comparison, even if it's apples and oranges. It seems to be part of our winning tradition. If you can
generate a score, than one can be compared to the other and therefore one can be better than the other.
This comparative process works as well as tabulating votes in Florida or counting gubernatorial
candidates in California.

There is also the sense that through assessment comes standardization. We have all heard the growing
concern that exists among some in authority that educational content is exercising too much freedom and
attending to too many diverse interests. By measuring all outcomes in a similar manner, there is little
room for the diversity that has been the cornerstone of our academic freedom. With the use of
standardized assessment modes comes standardized outcomes.

Placing so much of the institution’s perception of success on the non-instructional and non-curricular
outcomes prevents the university from completing its defined mission, which is to educate its students.
We are forced to position the institution and our units in a manner that guarantees success at the
conclusion of the assessment model not the education and intellect of our students. We are not only
teaching towards the assessments but also modifying the acceptance policies of students in order to place
the institution in the best light. In other words, marginal or remedial students, stay away. Taken to the
extreme, we are quickly moving towards a time when we will be accepting only students who meet
graduation requirements and assessment standards at the time of admission.
Currently in the State of Texas, those institutions preparing K-12 teachers are assessed based on the results of the EXCET examination. In order for the institution to remain in good standing, 70% of the students, as an aggregate and in all ethnic subgroups taking the EXCET examination, must pass. Institutions that take the more at-risk students become at-risk themselves if the students in any subgroup are not successful and do not achieve a 70% group success rate. Those institutions that fall below the 70% levels will lose their credentials to certify all teachers. Therefore, this assessment is promoting the acceptance of only those students who are most capable of passing the examination and leaving the rest by the wayside. To address this problem, some institutions have already become more selective in their admissions. Institutions that have traditionally accepted marginally prepared students are getting squeezed, so that in the foreseeable future these students will have no place to go. Is it any wonder that people are simply teaching to the test? This is quite a dangerous precedent, as the next step, which we currently are facing in the state, will be Cliff Notes and, through alternative certification methods, simply taking the test without any appropriate preparation.

As this session was to provide helpful hints and not just be a window on what is happening, I do have some suggestions on how to deal with the assessment shadow that hangs over us. With this session title, “Playing the Outcomes Game,” it seemed appropriate to pursue a sports metaphor, so if you will allow me:

The first suggestion is, when faced with a new assessment model be sure of the game you are playing. Don’t show up thinking you are playing basketball and those controlling the assessment are playing football. To be clear, know what is being assessed. We often enter the assessment process with a much better understanding of what we do than the person or group overseeing the activities. So, even if you think you can score more points dribbling and shooting, if they want you to kick, kick it. The tendency is to recognize the flaws in a process that has little or no focus, as they often do, and then proceed to correct the model so that the data will more accurately reflect what you think they want. You know how good your program is and you want to show that, even if it means supplying data that they did not ask for. My advice is to give them what they ask for and explain how this assessment does not accurately reflect the instruction provided in your discipline. I believe this is much wiser than revising the model for your discipline. If you give them different information, those in charge, who most likely did not have an understanding of the disciplinary differences to start with, will assume you are not complying with the requested material in order to hide something. You will wind up having to explain the data that you provided and still be required to supply them with what was originally requested, thereby having to develop two explanations rather than just one.

Second, know whose ball you are using in the game. When I was young, the guy who brought the ball was able to dictate some of the ground rules and set the game. Particularly when it is a new assessment, it is important that you understand the general premise of the game and clearly understand the consequences of the results prior to the commencement of the exercise. You do not want to be put in a position of “I didn’t know.” Knowing the rules at the beginning of the exercise allows you to ask questions and identify problems prior to the development of the requested data. It is at this time that questions are not perceived as disciplinary specific but rather clarifying general problems. As I said above, we are often manipulating the data rather than modifying the outcomes. If either becomes necessary, you have a greater opportunity to prepare or revise the base of data. For example, a former institution of mine was required to do an assessment of class size and credit hour production of instructors at the different academic ranks. The premise was that junior faculty produced more credits than senior faculty who received higher salaries and therefore were more expensive and less productive. The president of the university also wanted to prove that cost per credit in certain areas, a.k.a. the arts, was too expensive for the institution to continue. In my department with many full professors teaching primarily applied instruction, their loads and credit hour production were low compared to others who taught classes, even smaller sized classes. I was able to successfully argue that, by definition, applied private instruction was not a class, and therefore should
not be calculated into the totals. As a result, private lessons were excluded and our average class size and credit production were consistent with other disciplines across the campus. In this case, knowing the hidden agenda was most helpful.

Third, **make sure that you understand the rules of the game.** I don’t believe that you would put money down on a roulette table if you did not understand the game. Yet I hear from colleagues, that they plan to put, “anything” down to “be done” with the exercise. While you never know if these types of documents are read, for the most part, you can’t take the risk that they are not. Before completing the required documents, try to ascertain how the data will be used and who will see it. There is certain data about your unit that you share very cautiously with your peers. Therefore, knowing how the data will be used will give you the opportunity to present, or withhold, the most sensitive data in the most appropriate manner. With the example given above related to applied private instruction, that approach was successful because that data was only to be used for class size purposes. If it were also going to determine full-time faculty allocations, that approach would not be wise, as many full-time faculty were left out of the total.

Fourth, **understand what is required for a winning score.** You want to know not only what wins the game, but also what is needed to keep you in the game. Having this knowledge allows you to provide the appropriate amount of information. I am of the belief that it is not helpful to supply people with more information than is called for or that they can digest at any one time. That is, we develop a large amount of data about our students and our programs. Some of these data are related to accreditation and others are not. While I find this information useful, I am not motivated to share all of this information with others on campus. Therefore, it seems prudent to answer the questions and have additional information at hand if needed.

Fifth, **make sure you know who are the players on your team.** It is important to know who are the players on your side; this should include those in key administrative positions whom you may count as allies. This knowledge will help you determine who can be a partner in your success and whom you can turn to for support. Knowing this may also help you modify the information to suit a particular style of the person who will receive it. In addition, it is important to know if any other disciplines share your concerns. It stands to reason, if you have difficulty with an assessment model, others may also be in the same boat. With edicts from the legislature on an annual basis, it was not unusual for alliances to be formed between the arts and math, or the social science. This adds credibility to concerns if it extends over several disciplines.

Sixth, **when the play is broken up and there are no other options, improvise.** While this is an option, it is one that I would use most carefully. Most coaches will tell you that it is much safer to run a set play rather than improvise. However, if it appears that the information that is requested by the assessment presents the discipline in such a bad light then, as a last resort, I would revise the informational model. Providing different information answers different questions, which will require further explanation in the end. I strongly believe that there is a direct proportion between the amount of additional explanations that are needed to support your assessment and the perceived level of concern related to the results. The greater the concern, the more support material will be developed, the more supporting material, the greater the opportunity that some of it will be misunderstood.

Seventh, **it is important that at the post game interviews, you don’t let the other coach be the only one describing the game.** When the results of the assessment or accountability are brought forward, it is important that you explain any quantifiable results and what they mean in relation to your arts program. While others may take results to mean what they interpret, it is important for you to take the opportunity to interpret the numbers as you prepared and understand them. For example, if a legislator or upper administrator looks at the credit hour production of your studio faculty and compares it to that of an English teacher, it will be
important for you to explain the culture within the studio and how it relates to the credit hour production and the nature of the discipline.

Eighth, you don't have to win all of the heats to make it to the finals. This is important for two reasons. First, it does not always pay to be first across the finish line. Often, as a legislative driven assessment begins to develop, there is a realization that certain aspects won't work as planned and changes are made to modify the process or procedure. If you have completed your task too early, you might wind up doing it again and, as happens from time to time, the problems that you had with the process get solved before the due date. So in this case, first in may not always be the best. And second, sometimes artists need to fight the competitive nature to win at every opportunity. If you noticed during the Olympics, it was not necessary to win every heat to make it to the finals. This demonstrates that there are other ways to achieve success than simply being the best. Because our disciplines are different from others and because often, the benchmark of the assessment falls within another discipline, it is, sometimes, better to remain in the pack and stay more camouflaged. That way, you call less attention to your results and others are less likely to formulate challenges. Even if funding is a consequence, it is rare that all of your funding will be based on these assessments. Rather, you will need to weigh your response versus the available resources. I have been at institutions where other departments want to make applied studies an issue at every opportunity. There were many times we were the resultant few hundred dollars available through the assessment process was not worth the arguments that would follow from others bent on trying to make a string quartet meet a minimum enrollment of 12.

Ninth and last, if at all possible, wait as long as you can to see if the grounds crew can remove all of the rocks from the playing field to prevent some bad bounces. The ground is changing all of the time. As additional players become involved, I believe there is a greater chance for the positive aspects of assessment to come to the fore. Several of the regional accrediting agencies propose to evaluate the institution solely on its own mission. However, in this case the mission can vary to mirror the attainment of minimum goals and one is not sure whether the operations reflect the mission or the mission reflects the operation. Merely providing a quantitative assessment of the mission will not serve a positive purpose either. Both assessment models, when taken to the extreme, are greatly flawed, whether it is an assessment that has one base standard and compares everything against it or a completely moving standard that lacks foundation and substance.

The assessment pendulum continues to swing. The quality of educational outcomes is based on many different criteria ranging from faculty competence to facilities. We are finding that success is someplace in the middle. In the end, any worthy evaluation must recognize that each discipline is unique and that faculty must maintain ultimate control over the curriculum. Assessment of the quality of the many disciplines cannot be successfully quantified and compared just to find a simple utilitarian outcome, particularly for purposes that do not improve instruction. As an optimist, I believe the field of play is getting better and the bounces are truer. Little by little those advocates who understand the substance of the outcomes are assessing the outcomes. As in the Field of Dreams, we know, if we teach it, they will come. However, we want to be the ones determining if they are the real thing.