AN ASSESSMENT STRATEGY FOR ENHANCING SUSTAINED HKS EXCELLENCE

Richard Light
Harvard Kennedy School

HARVARD Kennedy School
SLATE | Strengthening Learning and Teaching Excellence

SLATE Working Paper 1:

An Assessment Strategy for Enhancing Sustained HKS Excellence

Richard J. Light

Contents

This Working Paper has three parts.

Part One: Why assess students' learning in courses and programs at the Kennedy School? What is new about this process? Haven't we been doing this for years? Don't students' grades tell us what we need to know about their actual learning?

Part Two: What are several different formats for assessing students' learning at HKS?

Part 3: What suggestions emerge for "institutionalizing" the assessment process, for constant improvement of future HKS courses and programs, based on lessons from SLATE projects over the past five years? We offer seven recommendations.

Part One: Why Do This? What Is New? Haven't We Been Doing Some Version of This "Assessment" for Years?

The Harvard Kennedy School has been evaluating its classes, and faculty classroom performance, for years. It has become routine at HKS, and indeed at many top universities, to collect on a regular basis several chunks of highly useful information from program participants. Each instructor or faculty member is rated by students. This gives Deans and Area Chairs at HKS valuable feedback on students' appreciation of each, individual instructor. In addition to feedback on the "quality of teaching" in each class, overall questions are asked about each course and its organization. This global feedback has been useful, was carefully put into place many years ago, and the quantitative summary of this data is now routinely made available to HKS students.

Recently the Kennedy School created SLATE – the acronym stands for "Strengthening, Learning and Teaching Excellence." In contrast to traditional students' rating of each class, SLATE's Assessment emphasis focuses on an entirely different question. We ask, "What are students actually, demonstrably, learning?" It is one thing to make a judgment about any individual students' "learning" while at HKS by looking at their official "transcript." A transcript from our registrar will reveal a string of courses taken and grades each student has achieved.

Suppose we look at a random student's transcript as they finish their first year in the HKS MPP Program. Suppose we find a string of all "A-"grades. Can we know from this string of grades what that student has actually learned in his or her year at HKS? Can we know what public challenges and analytic problems they might solve particularly well? Particularly poorly? SLATE believes the answer is no. To really understand more fully both what any particular student is truly learning and truly able to do, plus to understand what HKS is doing well

institutionally, and not doing quite so well, we need to "tap into" each student's capacity to solve actual public problems. This, in one sentence, is what drives our HKS Assessment projects.

What is new is SLATE's current effort to get a better handle, and more compelling evidence from students, about how each program actually affects both their measurable learning and their problem solving behavior. What did they learn, really? How do we know? Ultimately, our goal is to get useful information on how, if at all, each major HKS Program will help our students to maximize their effectiveness in their work.

The current focus began about five years ago. Several leaders on campus--including President Emeritus Derek Bok and Dean David Ellwood, with strong support from both Academic Deans and Administrative Deans--urged the SLATE team to get more compelling evidence about what our HKS students are actually learning. So it is in the spirit of these calls for systematic documentation of what students (in our HKS case - - both regular "Degree students" and also "Executive Education" program students) *are actually learning while at the university*.

Part Two: What Formats Are We Exploring to Better Understand the Actual Learning That Takes Place in Our HKS Programs?

We have used *FIVE different methods of exploring participants' learning*. Some have been tried more in depth than others. As we try each method, we are systematically learning which are most compelling, which are useful, and which are more questionable. It would have been difficult to predict from the outset all that we have learned. Here is a summary of the five formats we have tried out over the past five years, with a brief appraisal of each one.

Format One: Pre-program and post-program questions that invite qualitative responses

Using a modest number of open-ended questions, we have invited Executive Education participants to respond to these qualitative questions. We posed the questions for a formal "sitting" both when they first arrive for a program at Harvard and then again at the end of that program. The questions are intentionally worded almost identically "pre" and "post" program. We then look for changes in the responses. We have used this format in the Senior Executive Fellows Program three times, in the State and Local Government Program twice, and in the Senior Managers in Government Program once.

Examples of qualitative questions: "What do you consider the three most important elements of successful leadership in your organization?" "How might you overcome barriers to exercising successful leadership in your organization?" "What do you consider the most critical ingredients of helping teams to work together well, and to generate the most constructive possible ideas and solutions to tasks for your organization?"

These three questions illustrate items that participants have answered before programs began, and then again at the end of the program while still here at Harvard. For any Executive Education Program where a part of the curriculum is about exercising effective leadership, or where it was about organizing teams of people to get a job done well, we ideally would like to see the "post" program responses be more reflective and sophisticated than the "pre" program responses.

Notice that this format does not have any numerical "scale" for participants to fill out. Each question asks for a verbal response. *This means an external rater must read each response and assign it some sort of numerical rating*

so that we can combine, synthesize, and aggregate first our "pre" program findings, and then our "post" program findings, across all participants.

Format Two: Pre-program and post-program questions that invite quantitative responses

Certain questions can be posed by asking participants to simply rate themselves, either their own knowledge or their own performance, on a numerical scale. We used a 1 – 7 scale and we found such simple questions consistently useful. Clearly there is no absolutely "correct" scale to use. An argument could be made for using a 1 – 5 scale so that results from this "learning evaluation" are congruent in format with results from the traditional 1 – 5 scale that is used to evaluate HKS individual faculty members in Degree Programs, as well as the 1 – 5 scale used to assess overall aspects of each Executive Education Program.

An example to illustrate a simple quantitative self-rating: "On a 1 – 7 scale, where 7 is very effective and 1 is frankly ineffective, rate yourself specifically as a constructive member of a working group or team, that has been assembled to accomplish a task at your organization."

Using this simple, numerical format, and comparing each participant's pre-program self-described rating to their post-program self-described rating, we have a simple yet direct way of getting an insight about any changes that took place for a participant while he or she was here in an Executive Education Program. Of course the limitation of this format is that we are solely measuring the personal, self-judgments of students and participants in HKS programs. We are not tapping into any objective or external set of evidence.

Format Three: Pre-program scenarios and post-program scenarios that present a substantive question, usually a public problem, and ask each participant to write for twenty to thirty minutes offering a substantive response or solution for solving the public problem.

The substantive question is based on concepts and ideas that are covered in that particular Executive Education Program (or covered in an HKS Degree Program course). Of course we hope the post-program responses will be somewhat more rigorous and sophisticated than the pre-program responses, and incorporate some key concepts that were emphasized during the program or course.

When developing this format, the SLATE team informally describes this as "giving a final exam at the beginning and then again at the end of a course."

Who creates useful and sensible questions when SLATE uses this third assessment model? The questions for this format were developed based on topics that either a Degree Program's Faculty Chair or an Executive Program's Faculty Chair teaches and emphasizes during an actual group of classes. This way, the questions we use for evaluation of learning are in good alignment with a class or an entire program's goals and its curriculum.

An example to illustrate (from the State and Local Government Program):

You are the Mayor of a modest size city. The city owns some of the utilities, such as the water company. A large corporation comes to you and makes an offer to purchase the water company. They agree in writing to continue to provide water to the citizens in your city, for a fee. The amount they offer to pay is far higher than you ever imagined possible -- it is four times your gut level estimate of the value of the water company. Describe

briefly in approximately thirty minutes, the factors that you would consider in deciding whether or not to sell the city's water company to this corporation. Whatever your decision, how would you present your decision to the public?

Where did this question come from and why did SLATE use it? A version of this question was used in our assessment of learning that took place in the State and Local Government Program for seventy participants. Part of their two-weeks-long curricula involved privatization, part involved attention to how to value properties and enterprises run by a city or state, and part involved how to present decisions to the public, to citizens. This one, substantive question, created by the SLG faculty, enabled participants to demonstrate learning in all three of these areas. Our hope of course was that each respondent's post-program response would demonstrate clear learning, compared to that same respondent's pre-program response to the same question two weeks earlier.

Format Four: Follow-up telephone calls to a random sample of HKS former students (alumni of Degree Programs or Executive Education Programs), six months after they leave HKS and are "on the job."

The goal of these telephone calls is to get feedback on the personal and individual experiences of our alumni. We work hard to get responses that are not platitudes, but rather that focus on actual behaviors, or examples from their workplace, that they are willing to share with us. We ask them to reflect on their experience here, and to give examples of how what they learned here has specifically affected what they do at work. We press them for concrete details.

We are aware that a during a follow-up telephone interview, it is entirely possible for a participant to exaggerate, or even to lie, about his or her "post-program experiences." Yet we hope that the specific examples each participant gives will be compelling enough that they will suggest how, and indeed if, some genuine growth and learning took place because of that participant's experience in an HKS program.

Examples of telephone interview questions six months after an Executive Education Program (from the Kennedy School's longtime *Senior Executive Fellows Program):*

1. After you returned to work from the SEF program at HKS, how did you take advantage of your return to implement any specific changes that were genuinely triggered for you by participating in the SEF program?
2. Could you please share *one, very specific and concrete, example* of what you did, and then tell exactly how it relates to your SEF program experience?
3. What specific concepts that you learned at HKS, if any have proven most helpful to you in your work over the past six months? Could you share a specific concrete example of what happened at your work that illustrates how a concept from HKS was helpful?

Format Five: Ask each participant to write a short essay before the program begins, describing their specific, individual project, how they would implement them, how they would define a successful outcome, what barriers to successful implementation of their project they anticipate, and how they would deal with those barriers.

A typical essay might run three pages. It must be written by each participant sitting alone in a room just prior to our Executive Education Program. We certainly could implement a version of this "assessment" design in a single course in a HKS Degree Program.

Then, shortly before the end of the Executive Education Program, or a regular course, we ask participants to prepare a second essay. They take their same project, and they answer the same set of questions. Our hope yet again is, of course, that their planning and ideas about implementation and evaluation, and strategies for dealing with barriers to implementation will be noticeably more sophisticated, more thoughtful and rigorous. In other words, we hope the final, "post-program" essays will reflect ideas the participants were exposed to throughout the Executive Education Program or throughout a regular HKS course.

By comparing the quality of the "post" essays to the "pre" essays, we can get a reasonably direct comparison of how the same participant is conceiving of implementing his or her project in the field after they leave the Kennedy School. Grading should be done "blind" so that no grader knows if he or she is grading an essay written before or after a course.

This "pre" program and "post" program essay format could be implemented in a variety of courses. It could easily be adopted both in regular classes within Degree Programs, plus in several Executive Education Programs. It has one special advantage that other formats for assessing students' learning might not have -- this simple format can be used to demonstrate to students that they actually are learning. By simply sharing both the pre-course or pre-program essay plus the post-course or post-program essay that each student writes with the student, they can make their own comparison and see for themselves how much they have learned. Many students seem surprised at the extent of their improvement.

This is the format that was used by the HKS Executive Program for Pakistan Senior Civil Servants. This format resulted in our getting quite detailed "Pre" vs. "Post" comparisons, since we have them essay by essay, for each participant, in the Pakistan Program during spring of 2007.

The gain scores were valuable using this simple design, precisely because of what we learned. The gains in learning varied enormously among participants. We learned that approximately one-third of participants demonstrated enormous gains, one-third had modest gains, and one-third had basically no gains at all. This was useful information after the first year of this program, for adjusting the curriculum so the second year could generate more effective results. It illustrates the value of SLATE's using assessment to seek "sustained improvement" over a period of time.

Part Three

Now that we have "piloted" the assessment process on several individual HKS Degree Program courses, several individual Executive Education Programs, a cluster of MPP courses, and even the entire MPP Program at HKS, (which will be shared in detail in a forthcoming SLATE Working Paper), what are key takeaway lessons for the next several years? What have we learned about the most effective way to evaluate our students' real learning? What kinds of assessment data are most compelling when we strive to use such data for sustained improvement of our HKS curriculum and pedagogy?

Recommendations for Assessment's Next Steps

Recommendation 1

Whatever the details of plans for assessing students' learning in HKS courses and programs in the future, it is *far* too expensive and too time consuming to retain a regular, senior faculty member for a large fraction of his or her time, and to do this year after year. This means making the decision to organize an ongoing capacity to assess students' learning that is significantly cheaper, and sleeker. Most likely SLATE is a good vehicle for expanding this effort, capitalizing on its growing institutional experience. Any plan should also benefit from building on what we have learned so far, since we have made a substantial opening investment up front, both in putting forth effort and for gaining sustained experience with different assessment projects.

Recommendation 2

HKS has a yet untapped capacity, consisting of current faculty assistants and similar staff members who could (a) be trained to administer certain kinds of questionnaires, and (b) in some cases, to analyze the results. This clearly is not true for all HKS faculty assistants or staff; it may not even be true of a majority. Yet clearly some of these good people would love to participate in the learning assessment enterprise. Each would need to receive some training in how to analyze the data from learning questionnaires. This must be done thoughtfully, rigorously, and carefully. If it is done poorly, the entire enterprise will suffer a lot. If it can be done well, we view this as a great professional development opportunity. No new people need to be hired. No new expenses need to be incurred. The one expense would be to get a modest, piece of a faculty member's time who could teach selected HKS faculty assistants and staff how to interpret and analyze the "learning questionnaires" data. That person would need to carefully supervise the entire process, step by step, and that person would be the "guarantor" of quality control.

A first cut modest estimate of cost might be that summer support for one faculty member's summer, or for a significant fraction of a single summer salary for one faculty supervisor, should get the job done for the entire School. There would of course be a small cost in terms of time spent, to having faculty assistants and HKS staff members gather and analyze the data on students' learning under faculty supervision. This cost for HKS seems relatively modest, even small, when compared with what many other institutions do – they hire a team of several "institutional researchers," at a large multiple of the cost of training faculty assistants who are eager for exciting professional development. If faculty assistants genuinely feel that participating in assessment of students' learning is part of their professional development, and faculty leaders allow their staff to carve out some time to do this work, it can be a fulfilling format for everyone.

Recommendation 3

Using the "pre-program" and "post-program" model for gathering information turns out to be particularly compelling. Both individual HKS courses and Executive Education should continue to build more experience using this relatively simple learning assessment model. SLATE has not yet found a comparably compelling data collection plan. To use the statistical terminology for this plan, it uses each participant in a course or in any HKS program as "his or her own control." Using this simple model enables us to see change over time (if any) for each participant's

responses, and to document any changes. Using the pre vs. post comparison design enables us to share with each faculty chair, dean and program manager, the specific, verbatim changes in the responses from participants to tough, policy-oriented questions before versus after their class or program.

Recommendation 4

In future analyses of questionnaires to evaluate participants' learning, SLATE *should consider using samples* of pre – post comparisons. There is no compelling need to gather pre and post data from 100 percent of all the students, except when the "n" is relatively small. Each student could simultaneously be asked, indeed urged, to fill out his or her pre and post questionnaires, so that we have data for all students in a class or program. We then can choose a random, or a stratified, sample of respondents and do our analyses of learning on evidence from that sample.

The one circumstance where using 100 percent of all the data might be important is where findings from our modest size sample lead to genuinely troubling findings. In that case, we may wish to carefully dig into the full set of responses from all participants to explore in more depth whatever point is troubling, based on our analyses of sample data.

It's important to emphasize the value of using samples of participants' responses. One the reasons it took a longer amount of time than originally planned to analyze learning assessment data over the past few years, is that a few times we were swamped with mounds of detailed data. A modest size sample of responses, both pre and post, would have cut the analysis time substantially. This would have saved, cumulatively, months of time for analysis of results. Looking to the future, if HKS decides to expand the idea of assessing students' learning on a fairly regular basis, using sample data may reduce costs substantially.

Recommendation 5

Three formats seem especially useful for helping a Faculty Chair, or Program Manager, or a Dean, to really know how well our students are learning.

One helpful format is asking students' judgment about certain skills, or ways of thinking about how well they DO something (e.g. "work effectively in small teams to accomplish a task..."). In that case, asking for a numerical rating from each student, both pre and post any course or program, has two advantages. One advantage is that it is simply easy to tabulate. The second advantage is that the time it takes each participant to answer a group of well-developed questions is reduced substantially.

A second helpful format is to use actual, substantive scenarios, or project descriptions pre and post any course, from each student as our main data set to describe students' learning. It is one thing for a participant to simply assert they learned how to tackle a challenge or a problem that HKS leadership views as important. It is another thing entirely, and far more compelling, to ask each student to write a reasonably detailed and rigorous response to an actual, substantive question given to that person, both pre and post their HKS course or their Executive Education Program.

The key to making this second format work well is that the questions each participant is asked should be closely aligned with the content and goals of the course or program. This takes serious thought and planning, as it takes extreme care to develop good questions. For example, if a particular program explores the topic of privatization, a pre and post question about when and how to "privatize" a city's or state's asset might be especially appropriate.

A third useful format is to ask students to prepare their own essay, to solve a specific problem, both pre and post an HKS course or an Executive Education Program. For the Pakistan Program, to illustrate, each participant identified a specific "project" or "challenge" they needed to work on and run when they returned home to Pakistan after their several weeks at the Kennedy School.

Each person was asked to write an essay before starting our Executive Program, giving ideas for how to organize the project they say they would organize and run. Then, during the final week of the five-week program, they were asked to start over and to write a brand new, "final essay." Again the goal was to describe how to most effectively organize their own program. In our grading process we were able to identify, in concrete and powerful ways, various kinds of learning and skills development that had or had not taken place during the Executive Program.

All three of these useful options should be considered by each course that does learning assessments.

Recommendation 6

The idea of doing follow-up telephone interviews, with even just a modest size sample of participants, about six months after the end of a course or an Executive Program, is a highly productive, and low cost, way of gathering useful information. These follow-up telephone interviews were tried for two Executive Education Programs. The insights these calls generated often were specific and helpful. Having past participants give very precise and specific examples of how they used what they learned here, and how their learning helped them to make a decision they otherwise might not have made, provides powerful testimony as we try to understand the impact we have on our students.

Recommendation 7

The experiments in measuring students' learning that we have actually implemented over the past five years illustrate one specific problem that we MUST AVOID. In several cases, it was completely clear that students did not take this learning assessment seriously. This is especially true for a couple of the courses and programs during the post-program time period for completing questionnaires. Unfortunately, this "not taking the job seriously because our minds are practically out the door" has a very specific consequence. It makes our programs look worse than whatever positive impact they in fact are having.

To make this point specific, we had several participants in the SEF Executive Education Program who gave a great, thoughtful, verbal, two or three paragraph answer to a question on their "Pre" questionnaire. On a 1 – 7 scale, their answer was scored as a 6 or a 7. And then, four weeks later, that same participant scribbled an obviously rushed response, if indeed they gave any response at all, to the exact same question on the "post" questionnaire. As a result, the objective data from our graders show that this participant lost major ground as a

result of his or her participation in the HKS Executive Education Program. It is one thing if our participants are not learning enormous amounts –that certainly can sometimes happen, and when it happens, we need to do something about it. Yet the probability that a student or an Executive Education participant loses lots of ground during their time here, seems basically zero.

So the takeaway, which has been confirmed by several faculty members in private conversations, is that we need to get the students to take these pre and post-course learning assessments seriously, especially the post-program evaluations, where we have had some problems over the past five years. On one hand, we tell our students their responses to this "learning assessment process" are not a formal part of their course grades. Then on the other hand we have no way to "force" a student to take the assessment questions, post-course, seriously.

So far it has proved challenging to get students to treat "post-course questions" as seriously as we wish. This remains an ongoing challenge for SLATE. We have tried a certain degree of exhortation, of goodwill, and of conveying to students that their taking post-tests seriously can help "the next generation of students." Implementing this specific step well remains a work in progress.