CUFA BC

Professionalizing Academics*

A Submission from Andy Hira and Darcie Cohen

CUFA BC's 20 Questions for 2020 Project

October 22, 2009

*We would like to thank Simon Fraser University's Institute for the Study of Teaching and Learning in the Disciplines for their generous support of this research, and Adrienne Burke, Cheryl Amundsen, and Peter Kennedy of SFU for their very helpful comments.

Introduction

Higher education plays a vital role in society in terms of creating both good citizens and a capable workforce. Academics are thus key actors in socialization, adaptation, and self-reflection in society. It is no mistake that most academic enterprises are public enterprises, reflecting the collective goods nature of educational services. The triumvirate roles of teaching, research, and (public) service all dovetail in theory to serve together to accomplish these missions, including the internal mission of continual improvement. While of course there are many hard working colleagues producing good research and teaching, and providing valuable public service, an academic career is increasingly fraught with uncertainty, as a growing pool of PhDs finds career paths uncertain and tenuous. Moreover, it is common knowledge that the opacity of academic standards often lends itself to factionalization and politicization of departments, where control rather than excellence becomes the aim, as we can attest to from personal experience in various institutions and documented in decades of *Chronicle of Higher Education* articles. Perhaps of even greater concern is the lack of attention to creating and rewarding excellent teaching.

Despite recent efforts by professional societies to develop ethical guidelines for academics, they remain more in the realm of statements of principles rather than recommendations for accountable procedures (eg APSA 2008). It is no wonder, then, that faculty members report increasing stress and declining job satisfaction (Cooper and Nojima 2002, 168). As the pressure to publish continues to ramp up, there is a natural increase in the number of outlets, leading to even more pressure to specialize and more difficulty in keeping up with research. The proliferation of new publication outlets also leads to an emphasis on short-term quickly published research (Bruneau and Savage 2002, 84). The natural result is that the overall quality of research and the sharing of results across the fields (as all evaluation is still done by disciplinary departments) declines, as long-term research becomes reserved for the chosen few who can garner large grants. Policy impact becomes irrelevant, and new ideas (as opposed to adding more proof to existing literature) highly risky and poor investments of time.

There is even less time for those who do consider teaching and service to be equal obligations (though not proportionately rewarded) to act upon such convictions.

In the meantime, there is an increasing public call for some form of accountability in higher education. As Bruneau and Savage (2002, 58) state,

> A public university or college is, we assert, truly accountable when it ensures members of the public can readily survey teaching, learning, inquiry, and

> governance. It is accountable when open to scrutiny, open to participation, open to criticism, and demonstrably committed to critical research….

These authors (61) go on to note that the desire for information and accountability is what explains the popularity of *Macleans* ratings. Such ranking systems are statistically inaccurate, and fraught with "one size fits all" and other problems. They tend to reinforce already-resourced institutions, and somehow seem accurate with an occasional climber, often based on the definition of a particular criteria rather than a distinctive policy (Harvey Editor, 2008). It also makes possible that national research and teaching assessment exercises, as has been the case in the UK and Australia (Barrie and Ginns 2007), will eventually come to North America. We assert, therefore, that is incumbent upon higher education to seize this moment for self-reform, not only to avoid ill-informed pressures but also to improve our own performance. We trace these problems back to their original sources, namely the incentives for creating a surplus pool of Ph.D.s trained primarily for academic careers and the assessment systems that guide academic life. We believe that the lack of clarity in these systems is a large contributor to the denigration of academic life.

Success in any such struggle is unlikely in the short-term, as there are such strongly vested interests in the current system, interests that believe there is nothing wrong and that academic freedom carries these prices. But, as we argue below, ambiguity in academics that serves short-term interests comes only at the detriment of students, itinerant colleagues, and society at large, and is and will continue to increase indirect pressures to reform a troubled system.

What we are suggesting is academics simply be consistent with what they suggest should be done in the areas of good governance and research. By good governance, we mean commonly shared principles including: wise resource allocation, rule of law, equity, accountability, transparency, information flow and accessibility, the ability to plan for the long-term and measure progress, the ability for an organization to set collective goals and learn collectively, stable and clear duties and steps for promotion, and clear and stable rules for the succession of leaders; in other words, the basic principles of participatory democracy. By good research, we mean the principles taught in Methods 101, such as: clear conceptualization, identification of variables for measurement, validity, reliability, and accuracy of measurement, and coherent logic and empirical testing as part of an ongoing evaluation process, in other words, the very principles we use to evaluate research. If we applied these core principles to academic assessment systems, we would find ourselves more than capable of pointing to and growing the contribution of academics to society and the resources that are so vital to our necessary contribution to society.

Problems with the Way Graduate Programs are Run

There is widespread concern in the social sciences and humanities that we are producing too many PhD students. While job markets fluctuate, there has been an increasing panic by students who are increasingly being advised to open options outside of academia. The breakdown of the system is reflected in the augmenting exploitation of sessional, part-time, and graduate student teachers, developing into an academic underclass with limited prospects for competing with recent graduates from the high reputation universities. The reason for a continued and perhaps increasing glut of PhDs is a simple reflection of incentive structures. Departments without PhDs are not taken as seriously in research competition and thus have a much harder time securing grants. Most departments are rewarded based on the number of graduate students they have, and graduate students provide a very cheap labour pool for teaching. At the graduate level, this means taking on students that may or may not be suitable and/or may not be a good fit for the Department. All of which brings us to the fact that graduate supervision is often not rewarded.

There is, indeed, clear documentation that the number of part-time faculty members has ballooned in proportion to the number of full-time members in North America, reflecting the continually increasing number of graduates in all fields. Some estimates put the increase at 5X as fast as full-time members from 1970-2000. At the same time, the number of faculty members in the liberal arts has shrunk from 68% to just 56% as faculty hiring, reflecting enrollment trends, in the professions has increased. The unsurprising corollary is that the average age of full-time faculty has increased dramatically. Positive trends in terms of the rising proportion of women and minorities are also clear (Schuster and Finkelstein 2006, 40, 47, 58, 71 & 163). The increasingly difficult job market undoubtedly is in good part responsible for the escalating stress faculty members feel when they do get an academic post. For example, a mounting number of faculty members have previous non-academic job experience (Schuster and Finkelstein 2006, 99-103, 171)

"Casualization" of the academic labour market reflects not only over-supply, but the fact that there is a trend of using non-tenurable contracts instead as well as strictly lecture and adjunct positions (Schuster and Finkelstein 2006, 175, 232 and Louis 2009). In fact, there are now 3 possible "career" paths: tenure-track, renewable contracts, and non-renewable contracts (Gappa, Austin, and Trice 2007, 15). The end result is a vast spike in research and publication expectations prior to tenure (or for those still hoping for a post) and increasing service burdens as the proportion of full-time faculty members shrinks. Absent outside intervention, it seems unlikely universities could ignore the incentives to push graduate students out after a few years when their funding commitments run out, to make way for the next cohort of students (thus ignoring built up experience in teaching). These graduate students then become part of a growing labour force who string together part-time or sessional positions, with no benefits or consistent employment (Cross and Goldenberg 2007, 153-4).

There are other sources of the reported high levels of stress, disappointment, and alienation among young faculty members. There is a sense of arbitrariness in terms of the ambiguity of guidelines and secrecy of deliberations for hiring, tenure, and promotion as well as lack of consistency in terms of rotation of chairs and committees. Thus selected items, such as a handful of student evaluations, can be used strategically as weapons to sink them (Rice, Sorcinelli, and Austin 2000). Therefore, the need for having accountable hiring, tenure, and promotion systems is more important than ever.

Yet, most academic departments continue to allow the hiring process to be vulnerable to ideological and personal politics. One does not find in a general description for comparative politics that a significant minority want a game theorist or another faction a Marxist. Many think they should only hire candidates from departments with higher reputations. One only finds out about these things after the fact behind closed doors. Therefore part of the hiring process is discerning and navigating the winning "fit" for a coalition that will vote for the candidate. Sometimes this happens by consensus, where there clearly is a superior candidate who also fits. But in today's job market, it is not uncommon to see candidates with less publications, teaching experience, or real life work experience hired for political reasons (eg "least objectionable"). There is no way for a rejected candidate to know why from the form letter they came up short when they have a superior record of qualifications in the 3 areas of research, teaching, and service. Unlike other professions, there is therefore no learning curve from failure in many aspects of academics, including failure to find a position. Thus, while "fit" is certainly an important criterion, the lack of clear hiring procedures or results, and the fear of lawsuits preventing the adoption of such a system, open hiring up to abuse and continual factionalization of departments. It is no wonder therefore that, particularly in large research departments, there is a tendency to move towards a strong homogeneity of perspectives, just the opposite of what a lively intellectual life and strong teaching call for.

The oversupply of graduates based on the rigging of the academic market as explained below has meant that obtaining an academic position has become a privilege, rather than an expectation, for well-qualified graduate students, with exponentially increased research (publication and grant) expectations. While there are clear social benefits at the undergraduate level of a liberal arts education, it does not serve anyone's interest to produce a growing surplus of graduate students who will not find employment in their fields; as one colleague mentioned to me "getting an academic job these days is like winning the lottery." Yet any department ignoring the incentives to produce PhDs would be foolish and pay a heavy internal and external price. The perverse market incentives therefore call for greater information about labour markets and higher level intervention.

Occupational outcomes for graduate students need to be monitored by discipline over time and passed on to potential graduate students so that they can make informed decisions. Provincial and federal governments might need to limit the number of departments creating graduate students (particularly PhDs) and/or limiting the total number in order to restrict supply to more realistic levels. As Mann and Nelson Hochenedel (2006, 299), state, "This may seem heavy handed, but it's no worse than churning out 100 graduates in discipline X each year, *knowing* 75 of them have no real chance of ever gaining full-time employment in the field to which they will be dedicating up to ten years of their lives studying." (emphasis is theirs).

Tenure and Salary Review and Promotion

Tenure as Permanent Employment

The end of mandatory retirement exacerbates the growing surplus of PhDs and the likelihood of politicization leading to a lack of new ideas and innovation in academics. As is well known, academics developed differently as an occupation because of the need for academic freedom. The tenure system was designed to ensure intellectual freedom, not lifetime employment.

The tenure system highlights the growing binary nature of the academic labour market between the haves and have nots and stifles the possibilities for mobility between the shrinking tenured market and with non-academic markets. The tenured markets are becoming reserved for research stars, and non-tenured for those who teach under often difficult circumstances. Tenure is not necessary for academic freedom, and in fact further reinforces the notion of individual, rather than departmental/institutional achievement (Batterbury 2008). A number of institutions are experimenting with long-term contracts instead of tenure, such as the new Florida Gulf Coast University. Analysts suggest that contracts should be for at least 10 years to ensure adequate time for scholarly development (Pits and Kamery 2005, 76).

The perverse result of conflating tenure and permanent employment leads naturally to a sense of lack of accountability (whether justified or not) as there is less incentive for strong performance or improvement after achievement. Instead, the incentive system dictates vigilance to avoid scandals that would cause one to lose lifetime employment. Public opinion is strongly against continuation of tenure for this reason and post-tenure procedures have so far failed to satisfy basic criteria of accountability (Neal 2008). There is a sense of an intellectual aristocracy, one that has rights to do whatever they want and continue to get paid permanently, especially given the lack of tenure for a growing percentage of the professoriate.

Though tenure is supposed to guarantee freedom of expression, the reality is far different (Williams and Ceci 2007). In the current system, junior profs live in fear and are often forced to choose between factions or take their chances when their departments are divided. The notion of a decentralized university with independent departments whose members alone are solely qualified to evaluate their work also means that the levels of review and accountability, namely upper levels of administration, are usually reluctant to weigh in, unless there is an egregious case of clear sexual or racial discrimination. The only recourse for a member of a Department that has an individual or minority view is thus to appeal review in an ad hoc fashion to the Administration, risking the wrath of others within the Department. Though pressure declines after receiving tenure, there is no doubt that such systems nonetheless continue to create tensions at salary review and promotion and may well fail to respond to actual outcomes in any clear way in terms of teaching, research and service.

In a number of other systems outside of North America, such as Great Britain, tenure has been eliminated, while constructing stronger legal protections for job security in regard to freedom of academic enquiry and expression (Williams and Ceci 2007). It is worth pointing out that journalists and writers do not require such systems to speak out. Surveys of academics find instead a socialization towards timidity, undoubtedly reflecting the incentives of peer review (Ceci, Williams, and Mueller-Johnson 2006). Rather than being an exciting plane for intellectual creativity, then, the junior faculty members finds a stultified system of rewards for re-proving the findings of those considered prominent in their fields, thus reinforcing the supposed merit of such persons and institutions.

Tenure may allow for lower salaries in exchange for permanent employment and deep specialization, but it does considerably reduce institutional flexibility (Pits and Kamery 2005, 71 & 74). The overall lack of accountability in the academic process has led to rising public opinion against tenure, according to 2006-7 polls, one of which indicated that 82% believe it need to be changed. The perception of lack of accountability after tenure is great and perhaps justified, given the general lack of systems for post-tenure review, or, lack of post-tenure review entirely in many cases. Indeed a recent survey of Political Science chairs revealed that a substantial proportion believe that tenure has "shielded incompetent faculty members from dismissal" (Neal 2008).[1]

The ambiguity of tenure procedures and criteria in practice undoubtedly vastly multiplies stress for junior faculty. While potentially increasing flexibility, ambiguity also opens the way for personal and factional politics to enter into decisions, to be covered by claims such as those regarding quality of research. There appear to be few real efforts to resolve such problems. Duke, like other universities, had adopted a central administrative review

[1] The %s are: 49% for BA-granting institutions; 63% for MA; and 56% for PhD.

process for policies and procedures (Gappa, Austin, and Trice 2007, 201), however, it is likely that such efforts are more than paper exercises with little real measurement involved. Departments are reluctant to give any explicit thresholds for tenure and promotion. This gives them the leeway to get rid of "bad apples," uncollegial colleagues, but the sword of ambiguity again cuts both ways. In fact, it is precisely this ambiguity that leads candidates to stress, and where there is denial of tenure or promotion or unhappy salary review, to tie up resources in appeals and potential lawsuits. The problem must be particularly vexing for administrators who have to deal with appeals and possible legal action, yet have almost no input into the criteria or decisions themselves.

The composition of a tenure and/or promotion committee is therefore of grave concern when there are factions in departments. While an argument can be made that a regular rotation of the committee and clear procedures are vital, in the one-time instance of tenure, such precautions can in no way guarantee professionalism and objectivity. Deans are particularly reluctant to enter into the fray unless there is a clear lawsuit potential, as they tend to enshrine departmental autonomy based on pecuniary expertise of a field.

While we suspect that most tenure cases are probably done in good faith (as lawsuit potential pushes for greater diligence), the lack of standards causes an unnecessary level of uncertainty throughout academic careers. Even if one takes a rough measure of what senior colleagues have done, it may not be enough for tenure or promotion. If we believe tenure should continue, then we need to ensure that there is a greater sense of accountability for what happens at and after tenure in terms of job performance.

Assessment of Performance for Tenure, Promotion, and Salary Review

Assessment procedures for academic performance do not lend themselves to the basic principles of good research. The tenure process is a particular point of stress, related in good part to the "lack of comprehensive, clear, and rational guidelines and procedures (Gappa, Austin, and Trice 2007, 75)." Such pressures are naturally higher for women given additional duties of childcare, and for minorities who may feel excluded from the collegial club. The lack of studies of tenure and promotion processes (Rothgeb and Butler 2009) likely contribute to a lack of innovative reaction to the growing stress of faculty life. The lack of clear guidelines about what is expected from faculty members along with the oversupply of graduates has steadily ramped up pressure on faculty members, who never know when they have done enough.

This might seem surprising to an outsider, since almost all North American institutions explicitly lay out the criteria for tenure and promotion as adequate achievement in research, teaching, and service. In research-oriented universities, an "international reputation" in research is often evoked as the threshold for promotion to full professor. It is a well-known fact that most faculty believe that research counts for much more.

Analysts note that the proportion of faculty members who publish regularly rose from 1/2 to 2/3 from 1969 to 1998. Though there are still stark differences by institution type, liberal arts (4 year) universities have begun to increasingly emphasize publications and grants as well as excellent teaching throughout careers, as necessary for success (Schuster and Finkelstein 2006, 99-103, 171). The general tendency at most comprehensive institutions is simply to count the number of publications, thus eschewing quality as a factor (Youn and Price 2009, 218).

While lip service is paid in most cases to formal equality of research and teaching, and significance to service, the reality is that research counts for considerably more. Evaluation of teaching tends to be based on being "not cataclysmically bad," or just acceptable. Service is even more ambiguous, as it entails service to a variety of levels from the department to the university to the discipline and finally the general public. As the rewards for internal and general public service tend to be unrecognized in comparison to those to the discipline, where club benefits prevail (and international reputations ensue), there is a tendency to over-burden junior (recently tenured) faculty members with the internal chores (Macfarlane 2007, 267), thus reducing the sense of collective buy-in or mentoring by senior colleagues as well as the possibilities of using their experience to improve department performance.

Under such circumstances, external reviewers are one possible objective check. However, in world of dense specialization, the reality is that most reviewers will be have some connections with the candidate or his/her mentor, reflecting the Gramscian world of academics, based on social stratification as well as merit, thus biasing against new approaches or unconnected peers.

These problems are multiplied when we consider interdisciplinary scholarship (Hurtado and Sharkness 2008), given the iron chains of disciplinary organization of almost all North American universities, a system that militates against multi- or true inter-disciplinarity regardless of the amount of lip service paid to it. In Political Science surveys, 67% of PhD-granting departments in North America require articles in prestigious journals for tenure, while just 10% of BA and 19% of MA institutions do, (Rothgeb and Burger 2009) indicating that top notch interdisciplinary scholars are likely to end up in less prestigious universities, with less resources for scholarship. While most universities now proclaim interdisciplinarity on the one hand, they give all power and decision-making to individual departments on the other, revealing the severe blockages to cooperation across silos.

In fact, the lack of an open system and the extreme path dependency of academics are well-documented in studies (Youn 1988, 17). The system separates candidates into 2 discrete markets, one for teaching-intensive careers, and, for the elite, research careers (Youn and Zelterman 1988, 69). Reputation of institution, and networking with like-minded colleagues are major factors that determine who gets hired, published, and funded, and militate against new paradigms. Lest we be accused of hyperbole, we note the findings of a recent study of political science PhD hiring and placement, which concludes, "Comparing Figures 1 and 3 further reveals how remarkably little change has occurred in the centrality of the very top departments in the network over time, although some departments have become (marginally) more central and others (marginally) more peripheral, with only a few departments exhibiting substantial shift in relative location," (Fowler, Grofman and Masuoka 2007, 736), with location referring to the origin departments of hirees.

Peer review is also used for promotion and salary review. The peer review system for publication and promotion is also fraught with well-known problems. Absent clear guidelines, reviews of work can vary quite wildly; an author can feel that editors can cherry pick or over-determine end results based on their choice of reviewers. Such is the case particularly in the social sciences where there are often strong methodological divisions and perspectives. This leads to a natural clustering of scholars within camps for mutual protection, further entrenching divisions. A would be author frequently does not get clear feedback on what is wrong with the way he/she has approached the problem or examined it, preventing learning from taking place. The overwhelming nature of being an editor with the proliferation of production based on the job market probably leaves little time for helping a junior scholar to fix the problems with their research.

Internal review systems are even more problematic when (intra) disciplinary divisions exist. The inability to agree upon clear criteria for performance, ranging from quality standards for publication and teaching to what constitutes positive service, lead to an ambiguity that opens the way all too often for political gamesmanship. In terms of teaching, there is a tendency to rely on broad quantitative surveys of students at the end of term, which is a useful tool but does not really get at what students learned and what they do with that knowledge. In fact, most professors acknowledge that the standard for teaching in most comprehensive or research universities is simply to avoid too many complaints. "Collegiality" similarly can cut both ways. It can serve to ensure a smoothly functioning unit as well as to keep out women, minorities and new approaches (Adams 2006).[2]

Moving Forward on Tenure, Promotion and Salary Review Assessment Systems

How do we move towards a more transparent and accountable system for tenure? There appear to be no quick fixes. In the UK and Australia, a national academic evaluation system based primarily on number and place of publications has evoked considerable controversy. For example, highly ranked departments were closed down for lack of enrollments,

[2] Not surprisingly, studies show discrimination against women and minorities continues, though the situation has improved. See Cathy A. Trower, Amending Higher Education's Constitution, Academe, 94, 5 (2008): 16-18.

bringing into question whether the system of quality review had any consequences (Bruneau and Savage 2002, 106 & 117).

General guides to faculty review recommend a set of clear guidelines be given to junior faculty, and that performance reviews result in a career development plan that includes clear teaching, research, and service expectations and the steps that will be taken to achieve them for the next few (3-5) years. These should include the classes that will be taught and elements such as teaching workshops that are designed to strengthen capacity in each of the key areas. Faculty who are performing at a higher level should be rewarded at higher levels with larger salary increases as well as tenure and promotion. Informal rewards should also be used, including greater access to teaching and research assistants, differential assignments to areas of research/teaching strength, and travel and lab funds. However, success should also bring responsibilities for chairing important committees and mentoring junior faculty and graduate students (Curry 2006).

It seems to us setting these types of reasonable standards modified by type and mission of the university is the way forward, eg "we expect 4 peer-reviewed articles or 1 book for tenure or promotion," Such transparency would move mightily towards reducing tension and politicization, and allow us to concentrate on our central mission of research, teaching, and service. Having some transparency in terms of what counts and how much in regard to research output would also reduce considerably the common feeling that salary review is not a fair or open process. However, more fundamental to this question is how to assess teaching and service.

Teaching and Service Assessment

As academics, for almost all of us, our most profound influence on the world will be in training and shaping the next generation of students. Yet, Linda Nilson (2009) likens the challenge of teaching well under current circumstances as being asked to perform "magic." With increasing classroom size, ever more heterogeneous student populations, including large elements that are under-prepared, rising research demands, and higher level training needs and competition needed for employment on one hand, but shrinking resources on the other, teaching excellence is an increasingly difficult summit to climb. The most precarious aspect of that climb is the fact that we do not have any reliable system for assessing and rewarding teaching.

Therefore, a particularly worrisome outcome of the above trends and structures is that as research and service expectations ramp up, there will be less time and attention for teaching well. There is often a general perception that those who emphasize teaching are unable to conduct research on the same level. The fact that research is visible and brings in outside money further reinforces the notion that competence in teaching is all that is required (Serow 2000). Yet, there is no evidence that excellent researchers can not become excellent teachers (Hoffmann and Oreopoulos 2009, 92). In fact, the quality of the undergraduate education and the capacity of graduate students to teach should directly affect the reputation of the institution, the availability of researchers, and the intellectual vitality of the instructors, and thereby the ability to succeed in research.

Most academic units give teaching equal weight in their formal systems of review. Indeed, most Political Science departments in North America formally rely upon a combination of student evaluations, a teaching portfolio, a review of syllabi and peer reviews (Rothgeb and Burger 2009).[3] However, the reality is different. For example, in tenure and promotion decisions, teaching often does not figure into the external reviews. Peer review is not always carried out. There is almost no reward for attending teaching workshops (Cooper and Nojima 2002, 167).

Particularly problematic is the over-reliance on quantitative student surveys. Most universities simply look at the average for overall satisfaction, often ignoring other pertinent questions, such as amount of learning, level of workload, difficulty of course, and expected grade. More importantly, studies show distributions of grades have weak if any correlation with excellent teaching, and there is no chance to examine actual feedback given, or improvement in papers (Cooper and Nojima 2002, 167). In fact, researchers have found a wide variety of variables that directly affect evaluations, such as time of day, size of class, and level of class that are never taken into account. As Baldwin and Blattner state (2003, 29), "Most of us would not dream of giving our students a final exam as the only measure of their performance in the class." As Sadler (1999, 115) notes, students are "uncalibrated instruments." Such evaluations are likely to lead to selection bias in that the most unhappy people are the most likely to fill out surveys. It is no wonder that surveys show little faith among staff in student survey results, despite our heavy reliance upon them to gauge teaching (Douglas and Douglas 2006).

Quantitative student evaluations do not really get at the range of skills taught in a course. These might include "complex and transformative learning" that goes beyond facts towards developing thoughtful analysis, interpretation, and use (Wright no date). Course-based evaluations also do not get at a department's overall vision and goals for teaching (if such a thing exists!). As Van Note Chism (2007, 17) states,

> Specifically, experts indicate that while students are the most appropriate judges of day-to-day teacher behavior and attitudes in the classroom, they are not the most appropriate judges of the accuracy of course content, use of acceptable teaching strategies in the discipline and the like. For these kinds of judgments, peers are the most appropriate source so information… (including giving feedback on): subject

[3] It is quite interesting to note that between 69-72% of institutions state that they use peer review for teaching evaluation, depending on whether they are a BA, MA, or PhD-granting institution.

matter expertise, course goals, instructional materials and methods, assessment and grading practices, student achievement, professional and ethical behavior, and thesis supervision.

In short, we have not taken the basic methodological steps of considering the validity, reliability, and long-term outcomes of the evaluation for teaching performance. A recent in depth focus group study of teaching evaluations revealed that most student evaluations responded most directly to how the student "felt" about the course in terms of enjoyment level; whether the instructor was "likeable" and "not boring;" whether the professor agreed with the student's perspective on the subject matter; and whether a grade was given commensurate to their own perception of their effort. The author notes the fact that no follow-up studies of the methodological effectiveness of such surveys are apparent, that such scores are in fact used as the sole measure of teaching effectiveness, and that they are treated as ordinal (sliding scale) rather than the categorical variables that they are (as averages are compared) (Titus 2008). Basic statistical rules of validity, reliability, and tests of range and variance are largely ignored in such systems, though outcomes are considered as unimpeachable in review systems. Though more informative, national level surveys such as National Survey of Student Engagement in the US (http://nsse.iub.edu/index.cfm) simply replicate these limitations on a larger scale.

We do not even have a clear conceptualization of what teaching effectiveness means, but it is obviously multidimensional. Absent clear standards, criteria or feedback on teaching materials, deliveries, and outcomes, how could we possible expect teaching to improve in any systematic fashion? The end result is often the creation of a student as client/consumer culture in academics. While certainly students and parents need to have strong input into teaching, there are 2 obvious problems unless there are counter-balances. The first is the obvious tendency to grade inflation, as popular professors are rewarded for their ability to entertain, rather than to teach. The second is that students, many younger than 21, by definition are inexperienced and not experts in education, life, or the subject that they study, and thus under-qualified to be the sole judges of how to teach.

Peter Kennedy, a renowned international professor of Economics and winner of teaching awards at SFU, pointed out to us that what matters is the intellectual integrity and quality of the course; the integrity of the professor in fair evaluation of students and conveyance of the core material of the course; perseverance in continual improvement in knowledge and skills in both the area of studying and teaching itself; and courage, in risking disapproval by pushing students hard, taking an active public interest in teaching, and being willing to innovate in teaching. As Peter points out, education may likely take several years of lag time before its full effects becomes appreciated. Many of us have noted the value of the "hard" professors who pushed us, but only well after the fact.

Graduate education further perpetuates the cycle by centering around the production of top quality research, but with little to no attention in most cases paid to preparing faculty members for teaching or service (Mathews-Gardner et. al. 2008). Comprehensive studies of graduate education reveal a stultifying lack of preparation for academic life, with little systemization to graduate training in terms of expectations or guidance (Wulff, Austin, Nyquist, and Sprague, 2004). While some teaching is common as it makes sense for the system, there is generally no training for teaching. Moreover, it is unusual for PhD graduate students to formally learn how to orally present a conference paper, organize a workshop, or prepare a grant proposal, all items that will be part of their successful navigation of an academic career. This is all the more disconcerting when one considers the tightness of the job market, and studies that show that teaching effectiveness is formally considered as a key variable for hiring, even at large research universities (Meizlish and Kaplan 2008, 501). It is likewise interesting to note that there is no training for the increasingly time-consuming service tasks that occupy so much of our time and that have crucial effects on the quality of the education and research delivered.

Research on Teaching Rewards in Canadian Political Science

These concerns led us to find out what Political Science Departments across Canada are doing in terms of shoring up or supplementing these flawed systems. We note that our focus is on regular, full-time traditional (bricks and mortar) departments in Canada, while acknowledging the growing importance of distance and other education providers. While there is no way to adequately deal with each of these voluminous topics, we hope a review of the issues followed by some empirical discussion may move the conversation forward. We initially began our research with an on-line survey of practices, but found the need to supplement this with in person interviews due to poor response. We found that the issues we discuss in regard to teaching incentives are ones that most Canadian Political Science departments are struggling with.

During 2009, we therefore conducted an on-line survey with Canadian political science department Chairs. We sent out 55 surveys, receiving back 13 responses, for a response rate of 24%. We then conducted seven telephone interviews (using the same survey) to bring the total to 21, for a total response rate of 38%.

Table 1: Profile of survey respondents.

	Average	Range
Number of FT Faculty	14.6	3-35
Course Load	4.6	4 to 8
% buy-outs/year	27%	0-80%
Number FT Lecturers	1.4	0-5
Number Sessionals		
Has a Grad Program	71% yes	
Has a PhD Program	38% yes	

All responses have been anonymized, however, the responses reflect the diverse population of Canadian higher education, including responses from small, medium-sized (comprehensive) and large research universities. We note also that responses came from all over Canada, including institutions in British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, Québec, New Brunswick, Newfoundland and Nova Scotia. We would like to thank our generous colleagues for responding to the survey.

We asked a number of questions about how teaching is rewarded. Not all the questions were answered, or had limited response rates, nonetheless several interesting observations can be made. The first is that while formally salary review procedures put teaching on an even keel with research, *almost no one thinks teaching is given equal weight in reality*.

Table 2: How is teaching rewarded in salary review?

Salary Review - formal weights	Average	Range	18 responses in total, 3 n/a
% research	46%	0-70	
% teaching	37.60%	15-100	
% servce	21.40%	0-33	
Does formal weighting correspond to real?	72% say no		

We then asked a series of questions related to how teaching was evaluated. The responses are not that surprising. There is an almost exclusive reliance on student questionnaires. There is emphasis on adequate teaching at the tenure stage, but not on a periodic basis. There are more long-term studies of teaching according to our surveys than we would expect, so this response deserves follow-up study. Scrutiny of teaching appears to happen just once in many Departments, at the time of tenure.

Table 3: How is teaching generally evaluated?

How is Teaching Evaluated?			21 responses
Student Questionnaires	21		
Peer Review	6		
Professional Staff Eval.	2		
Teaching dossier	2		
Is Teaching Evaluated Differently in Tenure/promotion vs Salary review?			
	Yes	No	
	48%	19%	21 responses in total; 7 n/a, 10 yes, 4 no
Note: where peer assessment is used, it is only for tenure			
Is there long-term evaluation of teaching?, eg alumni surveys			
	Yes	No	
	29%	71%	21 responses

One particularly sticky question is how to handle workload equity. We found that most respondents use a regular rotation to take care of this problem, and about 40% consider workload during tenure and promotion.

Table 4: Teaching Equity

Teaching Equity - Large Course Rewards		20 responses
Rotation (taking turns)	9	
considered in tenure/promotion	6	
point system	3	
course release	1	
Other	1	

What efforts are political science departments in Canada making to improve teaching? Table 5 demonstrates that the main (and generally sole) avenue is through university awards. There is quite limited payback for attending workshops, despite universities putting considerable resources into such efforts. There is some effort to push PhDs to attend teaching workshops, but the results suggest that these are university-wide efforts, rather than political-science specific ones. New hires get course releases often, but no real training program for teaching political science.

Table 5: Teaching Improvement

Rewards for Attending Teaching Workshops?			
	Yes	No	
	10%	90%	21 responses in total
Teaching Awards?			
-University-level?	19	2	
-Faculty-level?	10	11	
-Departmental-level?	1	20	
Aid to Junior Scholars' Teaching			
	Yes	No	21 responses in total
Course Release in 1st Year?	17	4	
Research Stipend?	18	2	20 responses in total, 1 n/a
	Yes	No	
PhDs Required to attend Teaching Workshops?	66%	33%	6 responses in total, 15 n/a
PhDs matched w/faculty mentor?	37%	63%	8 responses in total, 13 n/a

New hires arrive at universities subject to a maelstrom of possible political agendas. They often suffer from any clear guidelines about what constitutes merit and the grounds for advancement in our profession. One interviewee suggests the adoption of a system of points given in ranges for different types of publications. The system does not deal with differences in quality of outlets (other than peer-reviewed), but it does at least give a clear rubric about how publications will be counted (eg articles vs. book reviews). The message new scholars get about teaching is clear- one interviewee put it bluntly, "you can't substitute teaching for scholarly activity and remain a credible university professor." Even traditional liberal arts programs have transformed themselves, partly through top down pressure, into research pressure cookers. One interviewee at such a place observed,

> "…this is an institution that is going through a metamorphosis. I think that traditionally this was an institution that really valued teaching and that was a really important criteria in terms of people's career advancement. There was a real emphasis on maintaining small class sizes, teaching was important in this institution but it has become rapidly less important because all of the criteria for promotion have to relate to research. Problems in this faculty include everyone trying to avoid teaching because they need to perform in other areas like career development. The problem with this is, of course, someone has to teach the students! There seems to be very little incentive to take teaching seriously at the moment."

Moving Towards a More Reliable Teaching and Service Assessment

There has to be some way to give appropriate signals (ie incentives) for assessing and improving performance. Obviously, the current system does not do so, leading to continual internal political battles over the "burden" of teaching and service. As one interviewee stated, "we used to have "merit." Lousy teaching and poor service may have prevented the odd person from getting merit, but very few people ever got merit for teaching or service. It has now been abolished in favour of a series of awards."

Graduate students need formal training for all aspects of their job if they intend to enter into academics. One initiative funded by the US government, the Preparing Future Faculty (PFF) program allowed for cross-institutional cooperation, including the ability of graduate students to meet faculty mentors from other universities, organized into clusters of different types of geographically close institutions (Gaff, 2007, 18-19).[4] Recreating a similar system in Canada would signal the importance of teaching and allow for a shared learning experience in terms of how best to train graduate students for teaching.

The evaluation of teaching also needs considerable renovation. Evaluation should be based on course materials, delivery, and outcomes (Fink 2002, 50). Departments should state the core topics, skill sets (in PS, these are often analysis, writing, research, and oral presentation), opportunities for real work experience, and a sequential set of courses, including possibly a capstone course, that allows students to build up and reflect upon the learning process (McClellan 2009, 48).

Generally, there are 2 approaches to achieving such goals on the departmental level. One is to set up a clear mission, evaluative criteria, assessment procedures and course planning from the top (university and then chair) down. The second one is for more decentralized situations or where there is some faculty disagreement, and allows the faculty themselves to get together and decide upon such matters. In either case, objective assessment, peer review and incentive systems are vital to success (Deardorff and Folger 2009). Once teaching matters, the process should be seen by faculty as a method of collectively improving their performance, recognizing the diverse talents of colleagues, and improving student outcomes. Assessment at the departmental or university level could include studies of alumni, current students, faculty, and employers. The wide variety of methods we use for our own research could be used here: from interviews to focus groups to surveys. For core materials, exams and placements could help to reveal collective success. Student evaluations are best done longitudinally (including after graduation), allowing time for the student to better put the class material into perspective and use (Kane, Williams, and Cappucccini-Ansfield 2008). These could include exit surveys of graduating

[4] http://www.preparing-faculty.org/

students and alumni asking about the performance of professors, techniques, and knowledge and skill bases.

It is extremely odd that we accept the problems of peer review with our research but do not embrace it as a tool for teaching. Hesitation to embrace it comes from our points about the politics of salary review. Supposedly it can only be done by someone with disciplinary knowledge (a point with which we disagree), therefore it opens up a potential conflict of interest and personal politics to a fellow member of the department. Even if it's voluntary and not counted, there are likely to be strong suspicions by some members of the department that such information could be used against them in salary review or promotion. The obvious solution would be to have an outside person, perhaps someone from a neighboring university, do the review. This brings into play a resources and coordination problem, thus, such a system could work only occasionally, perhaps every few years. For young scholars who need intensive guidance, pairing with a mentor would make the most sense. On the other hand, rather than being latent resources that are underutilized such as poorly attended teaching workshops, teaching staff at the university level could be an integral part of improving instruction and spreading knowledge across universities, if they were a required part of the assessment of faculty teaching.

On a class level, a pre- and post-test would reveal improvement in student outcomes, rather than simply how they feel at the moment of angst (generally right before a final exam) about the experience. A recent article (Stark-Wroblewski, Ahlering and Brill 2007, 410), conducted such an experiment recently in a psychology course. They found, reinforcing the conclusions of other literature, that there was quite a low correlation between improvement in test scores and student evaluation scores – just 0.18. They note the slightly higher correlation, however, with the specific question of whether the student learned in the course, 0.20. Such tests could be a valuable instrument to help move the speculative nature of teaching results forward. Yet, even such arrangements require supplemental peer review observation, to pick up the more subtle aspects such as learning to analyse, interact with peers, and make decisions in complex environments. Others note that such tests are inadequate for evaluating real learning as opposed to reproduction. They suggest other instruments, such as a student journal to be kept during the class with observations (Voparil 2009, 27), as well as the peer and expert review mechanisms discussed above.

Macfarlane (2007, 270) offers some suggestions about (internal) service. He suggests that a service portfolio can be created, with attention to criteria, such as engaging audiences; mentoring; organizing and representing; reviewing; and advising. More fundamentally, we need to come out of our ivory tower once in a while to reveal the contribution we are making and interact with the public so that our work is more recognized, valued, and responsive.

Conclusion

What we have seen above is that academics systems are poorly designed from the point of view of evaluation of results and accountability and transparency of systems. The basic management literature suggests that performance stems from the development of clear strategic targets and priorities; appropriate administrative and incentive structures to support, enhance, and evaluate performance; and the development of collective values towards continual improvement on an institutional, not just individual, level (Wang 2010, 10). Instead, the system incentivizes individual, not collective gain. Individuals are rewarded primarily for spending their time racking up publications. Teaching (and service), by contrast, is more inherently collective than often acknowledged (Fairweather 1996, 205). A proper teaching program requires coordination across colleagues for a coherent training program for graduate students; a clear step-by-step course sequence organized around knowledge accumulation; a sharing of experiences and knowledge to improve departmental and university-wide performance; and a series of objective tracking measures to gauge performance longitudinally. In fact, it is hard to begin to describe how well we are doing as there is no baseline for measurement.

We need to get back to recognition that universities are public institutions and therefore have to have demonstrable voice, transparency, and accountability mechanisms. It is time to get back to the original notion of service, which is for universities as public institutions to provide clear social goods. It is the dissonance between the public expectation that this is what taxpayer dollars are funding in paying for research and the peculiarities of the evolution of academic governance towards disciplinary-, small n-peer-reviewed aims that creates the fundamental problem. The general rhetoric of interdisciplinarity but the lack of follow-through in creation of such programs highlights the lack of real responsiveness in academic governance systems. The exceptional success of several new interdisciplinary programs in Canada in globalization, development, environment and international studies highlight that many parents and students see such programs as more relevant to today's problems.

More fundamentally, we need to re-orient faculty members towards a scholarship of teaching and learning as well as the particular niche of their research. If peer review is the way we judge scholarship in research, we should use the same procedures to evaluate teaching. This could be done within disciplinary norms if need be. An occasional external blind peer review of a teaching portfolio including an in class visit would be the optimal way of getting feedback on teaching (Steves and Rosegard 2009). Similar to current external reviews of departments (or as part of them), it could be done as a group report by a committee of the external assessor, a department member and a member of the university in a different department. This would reduce bias, and, when done by people recognized for teaching capability within the discipline, lead to an optimal feedback loop in terms of teaching improvement. The focus of the review could be on how to improve and fix problems, rather than the current emotion-laden diatribes that are embedded in student comments, the only source of feedback in most academic units. Since academic culture is presently biased against forms of discipline, the results of the reviews could be the possibility of a bonus for excellent teaching based on the reviewers' findings.

Absent greater attention to balance, the scenario for future academics is more bleak than the present situation. What is likely to happen is greater separation between researchers and teachers, and tenured and non-tenured faculty. Researchers will be those who can pull in external grants, with strong club effects. However, if they are tenure track they will also be overburdened with internal service, administrative, and mentorship duties as their proportional number shrinks. Those who are excellent teachers but poorer (but competent) researchers will find decreasing recognition, if they are able to get onto the tenure track at all. One interviewee stated that their unit puts those who are primarily teachers onto a different track, with higher teaching loads but also greater weight given to teaching. This would presume the development of clearer teaching assessment mechanisms.

This may work in some selected cases, but it is hard to see in either scenario that students or society will generally be better served by separation of the duties. While there should be some exclusive places for top notch researchers and teachers, the majority of the academy needs to be able to do both, particularly at the level of graduate training where exposure to research is a key ingredient. Otherwise, we could end up with the scenario now occurring in some natural sciences where graduate students are never really taught to become independent scholars; they continue as low-level researchers until their funding is cut off and then struggle to find suitable employment. The key challenge then, is to find acceptable measurement systems for teaching and service and to adjust the practice of tenure, promotion, and salary review accordingly to create new, accountable, and transparent systems based on merit, one that engenders a culture that embraces excellence in teaching.

References

Adams, Mark L. 2006. The Quest for Tenure: Job Security and Academic Freedom. *Catholic University Law Review*. 56: 67-95.

American Political Science Association (APSA). 2008. *A Guide to Professional Ethics in Political Science*. 2nd ed. Washington: APSA.

Baldwin, Tamara and Nancy Blattner. 2003. Guarding against Potential Bias in Student Evaluations: What Every Faculty Member Needs to Know. *College Teaching*. 51,1(Winter): 27-32.

Barrie, Simon and Paul Ginns. 2007. The Linking of National Teaching Performance Indicators to Improvements in Teaching and Learning in Classrooms. *Quality in Higher Education*. 13,3: 275-86.

Batterbury, Simon. 2008. Tenure or Permanent Contracts in North American Higher Education? A Critical Assessment. *Policy Futures in Education*. 6(3): 286-97.

Bruneau, William and Donald C. Savage. 2002. *Counting Out the Scholars: How Performance Indicators Undermine Universities and Colleges*. Toronto: James Lorimer & Co.

Ceci, Stephen J., Wendy M. Williams, and Katrin Mueller-Johnson. 2006. Is tenure justified? An experimental study of faculty beliefs about tenure, promotion, and academic freedom. *Behavioral and Brain Sciences*. 29: 553-94.

Cooper, Joanne E. and Sheryl E. Nojima. 2002. Teaching: Academic Whitewater Rafting. 163-178 in Cooper and Dannelle D. Stevens. *Tenure in the Sacred Grove: Issues and Strategies for Women and Minority Faculty*. Albany: State University of New York Press.

Cross, John G. and Edie N. Goldenberg. 2007. Off-Track Vetting. 144-54 in Leigh Deneef and Craufurd D. Goodwin. *The Academic's Handbook*. 3rd ed. Durham: Duke U. Press.

Curry, Theodore H. 2006. Faculty Performance Reviews. *Effective Practices for Academic Leaders*. 1,2(Feb): 1-16.

Deardorff, Michelle D. and Paul J. Folger, Making Assessment Matter: Structuring Assessment, Transforming Departments, 77-96 in Michelle D. Deardorff, Kerstin Hamann, and John Ishiyama, eds. *Assessemnt in Political Science*. Washington: American Political Science Association.

Douglas, Jacqueline and Alex Douglas. 2006. Evaluating Teaching Quality. *Quality in Higher Education*. 12, 1: 3-13.

Fairweather, James S. 1996. *Faculty Work and Trust: Restoring the Value of Teaching and Public Service in American Academic Life.* Toronto: Allyn and Bacon.

Fink, L. Dee. 2002. Improving the Evaluation of College Teaching 46-58 in Kay Herr Gillespie, Linda R. Hilsen and Emily C. Wadsworth, eds. San Francisco: Jossey-Bass.

Fowler, James H., Bernard Grofman,a nd Natalie Masuoka. 2007. Social Networks in Political Science: Hiring and Placement of Ph.D.s, 1960-2002. *PS: Political Science and Politics*. 40,4: 729-39.

Louis, Deborah. 2009. Adjuncts: Solutions for a Mistreated Majority. *Chronicle of Higher Education*. 55 (June 12): no pages

Gaff, Jerry G. 2007. Faculty in the Variety of American Colleges and Universities, 11-22 in A. Leigh Deneef and Craufurd D. Goodwin. *The Academic's Handbook*. 3rd ed. Durham: Duke U. Press.

Harvey Editor, Lee. 2008. Rankings of Higher Education Institutions: A Critical Review. *Quality in Higher Education*. 14:3: 187-207.

Hoffmann, Florian and Philip Oreopoulos. 2009. Professor Qualities and Student Achievement. *The Review of Economics and Statistics*. 91(1): 83-92.

Hurtado, Sylvia and Jessica Sharkness. 2008. Scholarship is Changing, and so Must Tenure Review. *Academe* 94, 5 (no pages).

Kane, David, James Williams and Gillian Cappuccini-Ansfield. 2008. Student Satisfaction Surveys: The Value in Taking an Historical Perspective. Q*uality in Higher Education*. 14,2: 135-55.

Macfarlane, Bruce. 2007. Defining and Rewarding Academic Citizenship: The implications for university promotions policy. *Journal of Higher Education Policy and Management*. 29, 3 (Nov): 261-73.

Mann, Douglas and Heidi Nelson Hochenedel. 2006. A Manifesto of the Twenty-First Century Academic Proletariat in North America, 291-306 in Robin Barrow and Patrick Keeney, *Academic Ethics.* Burlington, VT: Ashgate.

Mathews-Gardner, Lanethea, Michelle D. Deardorff, Grant Reeher, William Hudson, MaryAnne Borrelli, and Glen Halva-Neubauer. 2008. Getting a Job at a Teaching Institution- and Then Succeeding: A Q&A with Experienced Teacher-Scholars. *PS: Political Science and Politics*. 41,3: 575-8.

McClellan, E. Fletcher. An Overview of the Assessment Movement. 39-60 in Michelle D. Deardorff, Kerstin Hamann, and John Ishiyama, eds. *Assessement in Political Science.* Washington: American Political Science Association.

Meizlish, Deborah and Matthew Kaplan. 2008. Valuing and Evaluating Teaching in Academic Hiring: A Multidisciplinary, Cross-Institutional Study. *The Journal of Higher Education*. 79,5: 489-512.

Neal, Anne D. 2008. Reviewing the Post-Tenure Review. *Academe.* 94,5 (Sept/Oct): (no pages).

Nilson, Linda B. 2009. Editor's Introduction: The Educational Developer as Magician, 3-13 in Linda B. Nilson and Judith E. Miller, *To Improve the Academy: Resources for Faculty, Instructional, and Organizational Development*, vol. 27. Somerset, NJ: Jossey-Bass.

Pitts, Sarah T. and Rob H. Kamery. 2005. Academic Model or Corporate Model: Is Tenure Still Viable? *Academy of Educational Leadership Journal.* 9,1: 67-79.

Rice, R. Eugene, Mary Deane Sorcinelli, and Ann E. Austin. 2000. *Heeding New Voices: Academic Careers for a New Generation*. Washington: American Association for Higher Education.

Rothgeb, John M. and Betsy Burger. 2009. Tenure Standards in Political Science Departments: Results from a Survey of Department Chairs. *PS: Political Science and Politics.* 42:513-19.

Sadler, D. Royce. 1999. *Managing Your Academic Career: Strategies for Success*. St. Leonards, Australia: Allen & Unwin.

Schuster, Jack H. and Martin J. Finkelstein. 2006. The American Faculty: *The Restructuring of Academic Work and Careers.* Baltimore: The Johns Hopkins University Press.

Serow, Robert C. 2000. Research and teaching at a research university. *Higher Education*. 40:449-463.

Stark-Wroblewski, Kimberly, Robert F. Ahlering and Flannery M. Brill. 2007. Toward a more comprehensive approach to evaluating teaching effectiveness: supplementing student evaluations of teaching with pre-post learning measures. *Assessment & Evaluation in Higher Education*. 32,4 (Aug): 403-15.

Stevens, Cheryl A. and Erik Rosegard. 2009. Establishing External, Blind Peer Review of Scholarship of Teaching and Learning Within the Disciplines, 311-331 in Linda B. Nilson and Judith E. Miller, *To Improve the Academy: Resources for Faculty, Instructional, and Organizational Development*, vol. 27. Somerset, NJ: Jossey-Bass.

Titus, Jordan J. 2008. Student Ratings in a Consumerist Academy: Leveraging Pedagogical Control and Authority. *Sociological Perspectives.* 51,2: 397-422.

Van Note Chism, Nancy. 2007. *Peer Review of Teaching: A Sourcebook*. 2nd ed. Bolton, MA: Anker Publishing.

Voparil, Christopher J. 2009. Assessing for Understanding: Toward a Theory of Assessment as Learning. 17-38 in Michelle D. Deardorff, Kerstin Hamann, and John Ishiyama, eds. *Assessment in Political Science*. Washington: American Political Science Association.

Wang, Xiaohu. 2010. *Performance Analysis for Public and Nonprofit Organizations.* Toronto: Jones and Bartlett.

Williams, Wendy M. and Stephen J. Ceci. 2007. Does Tenure Really Work? *Chronicle of Higher Education*. 53, 27: B16.

Wright, Barbara D. no date. Guidelines for Good Assessment of Student Learning at the Department or Program Level. Working paper, found at http://apsanet.org/content_13026.cfm

Wulff, Donald H., Ann E. Austin, Jody D. Nyquist, and Jo Sprague. 2004. The Development of Graduate Students as Teaching Scholars: A Four-Year Longitudinal Study. 46-73 in

Wulff, Austin & Associates, *Paths to the Professoriate: Strategies for Enriching the Preparation of Future Faculty*. San Francisco: Jossey-Bass.

Youn, Ted I.K. 1988. Studies of Academic Markets and Careers: An Historical Review. 8-27 in David W. Breneman and Ted I.K. Youn, *Academic Labor Markets and Careers*. Philadelphia: The Falmer Press.

Youn, Ted I. K. and Tanya M. Price. 2009. Learning from the Experience of Others: The Evolution of Faculty Tenure and Promotion Rules in Comprehensive Institutions. *The Journal of Higher Education*. 80,2 (March/April): 204-37.

Youn, Ted I. K. and Daniel Zeleterman. 1988. Institutional Career Mobility in Academia in David W. Breneman and Ted I.K. Youn, *Academic Labor Markets and Careers*. Philadelphia: The Falmer Press.

Appendix- Ethics Statements

My name is Andy Hira and I am an Associate Professor of Political Science at Simon Fraser University. I am writing to ask for your cooperation in conducting a short interview with either me or Darcie Cohen, my research assistant. The survey is being funded by SFU's Institute for the Study of Teaching and Learning in the Disciplines. The purpose of the study is to find out how other Departments of Political Science recognize and reward teaching. As you know, this is a common problem. How to give teaching its fair weight among research and service in evaluations of performance; how to evaluate teaching beyond student questionnaires, and how to train PhDs for future teaching, are all unclear issues. Our hope is to learn from others about what the current state of practices is, and to synthesize new ideas to construct better practices. We think we can make an important contribution that could benefit all universities facing these issues. We intend to publicize our results in future CPSA meetings.

We believe that such information is already public in the sense that every university knows its Department procedures, so we are just synthesizing that information. Nonetheless, I would like to emphasize that the survey is voluntary and that your input will be anonymized. If you have any doubts or questions, please contact me at ahira@sfu.ca If you have any questions regarding ethics procedures, you may contact Dr. Hal Weinberg, Director, Office of Research Ethics at hal_weinberg@sfu.ca or (778) 782-6593

Bios

Andy Hira is Associate Professor of Political Science at Simon Fraser University. He specializes in technology and energy policy, and is the author of *An East Asian Model for Latin America?* (Ashgate 2007). Andy can be contacted at ahira@sfu.ca

Darcie Cohen is a second year Masters student in the Department of Political Science at Simon Fraser University. Her research interests include electoral systems, green parties, energy and environmental policy.