



The Jackson Hole
Higher Education Group, Inc
2060 N. Spirit Dance Road
P.O. Box 9849
Jackson, WY 83002-9849
Phone: 307.734.1207
Fax: 307.734.0249
Email: jhheg@aol.com
EIN: 83-0317892

Evaluation of the Ontario College Quality Assurance Services PQAPA Pilot

William F. Massy
December 1, 2006

Pursuant to a request from the Ontario College Quality Assurance Services, I have evaluated the Service's PQAPA Pilot Project. My review consisted of examining documents and discussing the project with auditors, auditees, administrators, and OCQAS Management Board members during a visit to Toronto on November 12-15, 2006. I have tested the results against my understanding of global best practice in higher education quality assurance—an understanding that is based on participation in and study of QA processes in the USA, UK, Hong Kong, Singapore, Australia, Denmark, Sweden, and other countries.

My overall conclusion is that PQAPA meets its stated goals, is consistent with global best practice, and is the right process for Ontario. I'll elaborate these conclusions and offer suggestions for improvement under four rubrics: (i) audit as the approach of choice; (ii) PQAPAs guiding principles; (iii) the audit methodology, and (iv) the quality assurance and improvement criteria.

Audit as the Approach of Choice

The PQAPA *Orientation Manual* points out the distinction between "Quality Assessment" (the mechanism or procedures used to determine the extent to which quality exists) and Quality Assurance (the mechanism of procedures used to assure or measure the level or existence of quality)." The distinction is well accepted around the world, though jurisdictions differ on their choice of approach. Agencies choosing Quality Assessment evaluate quality levels themselves—for example through site visits and employer surveys as in Denmark and the system of assessment against subject-level detailed performance benchmarks in effect in the UK until 2002. Agencies choosing Quality Assurance evaluate the mechanisms and procedures *used by the institutions* to assure and measure, and also to improve, the level of quality—a process that has come to be called "academic audit." PQAPA falls squarely in the audit category.

There are significant reasons for preferring audit to assessment and I believe global practice is trending in that direction. Assessment's one big attraction is that it purports to provide externally-generated objective evidence about teaching and learning quality. However, evidence is accumulating that this promise is a chimera. The weaknesses of assessment are as follows.

- Assessment is adversarial and intrusive. It substitutes the site visit team's views about quality for those of the local academics, which disempowers the people on the ground and risks making quality assessment "someone else's responsibility". It also motivates institutions and their faculty to withhold information about shortfalls wherever possible. Hence it's virtually impossible to combine assessment with improvement in an assessment regimen.
- Valid assessments of the delivered quality of teaching and learning are difficult or impossible during a short site visit, even when supplemented with surveys or external examination results. Postsecondary education is a complex enterprise whose quality can only be understood by "living in the system" for a considerable period of time. For example, even surprise visits to a few classes can't tell the team very much about the institution's teaching and learning environment.
- A comprehensive assessment program is expensive. Assessment must be conducted at the subject level by teams made up of experts from the discipline being reviewed or closely allied ones. This requires $N_i \times N_s$ separate site visits (as well as the self-studies leading up to them), where N_i is the number of institutions and N_s is the average number of subjects per institution. And because one cannot use assessors with ties to a given institution, one often must go far a-field to populate the site visit teams—which further adds to expense.

Academic audit avoids or mitigates these difficulties. It views the responsibility for quality as residing unequivocally with the institution. The auditors' job is to ensure that the responsibility is being discharged effectively, a task that *can* be accomplished in a short site visit. For example, faculty who don't regularly participate in quality assurance and improvement find it impossible to discuss these matters coherently—whereas nearly all faculty can "tell a good story" about quality itself. The cost of audit scales proportionately to the number of institutions rather than institutions times subjects. Furthermore, because specific disciplinary knowledge turns out to be not so important in evaluating QA methods, the cost of populating the teams is lower than for assessment.

The core principle of audit is that better quality within the institution pays off in better education. And because the auditors evaluate the institution's quality measurements, the ability to document quality will get better over time. This is consistent with the "Minister's Binding Policy Directive," cited on page 6 of the *Guidance Note*: that each college has the overall responsibility for quality and that it should "incorporate processes that 'ensure the ownership of quality'." The job of audit is to help the colleges meet this responsibility and assure the public that they are doing so—in other words, "trust but check".

The UK ran both an audit and an assessment system until the formation of its academic Quality Assurance Agency (QAA) in the mid-nineties. The QAA tried initially to standardize on assessment but found it unworkable. It has since adopted audit as its primary QA method, with assessment being limited to institutions that have "failed" their audits. Hong Kong's University Grants Committee evaluated the merits of quality assurance v. assessment a decade ago and came to the same conclusion as the designers of PQAPA. Its two audit exercises proved successful, as have the programs in Australia, New Zealand, and other countries. Denmark is

experimenting with audit and it has been introduced successfully by the University of Missouri System and the Tennessee Board of Regents. An increasing number of U.S. accreditation agencies are using one or another variant of audit.

Guiding Principles

The *Orientation Manual* sets forth eleven principles, approved by the Management Board in September 2005, for guiding the development of the PQAPA. I believe the principles are appropriate and that PQAPA broadly adheres to them as indicated below.

- a. “Be consistent with best international practice.” Placing this goal in first position is itself a reflection of best-practice quality work. The thrust of this report is that, except in the relatively few areas singled out for improvement, it has been met.
- b. “Be effective, efficient, and cyclical.” PQAPA appears to have been conducted in an effective and efficient way—especially considering that much had to be invented because this was a pilot round. As pointed out, the process needs to be cyclical in order to achieve the hoped-for results. It’s too soon to say whether that will be achieved but the prognosis appears favorable.
- c. “Give appropriate notice to colleges prior to review.” The notice may have been sufficient but, as noted below, the orientation could have been better. I recognize the difficulties posed by last spring’s labor action, however.
- d. “Be described and operate as an audit...” PQAPA falls squarely in the audit category. Its documentation and the reports that have been produced bear this out.
- e. “Use and external, competent, and mutually agreed-upon objective peer review panel.” The panels appear to have been objective, clear of conflicts of interest and, based on my personal observations and report reading, highly competent. I found no evidence to indicate they were not mutually agreed-upon.
- f. “Follow a mandatory and clear process that identifies strengths and weaknesses, recommends improvements, and facilitates positive change.” Like all well-run audits, PQAPA identified strengths and weaknesses and based on testimony from the auditees, it facilitated positive change. The “mandatory and clear” dictum was not fully achieved, but that is not surprising given the necessity of experimentation during a pilot. However, I suggest that best practices achieved during the pilot be agreed upon and recommended strongly to future audit teams.
- g. “Distribute the preliminary report to the colleges for their comment for their response and include the college response in the final report.” The principle represents an important element of best practice and it appears the preliminary report has in fact been so distributed. Only one of the three draft reports I received includes the colleges’ comments, but this may have reflect the fact that they were works in process.
- h. “Provide for categories of approval.” Such categories have been provided for the individual quality criteria but not for the audit as a whole. I’ll offer a suggestion about that in due course.

- i. “Provide for a method and system of appeals against its decisions and reports.” My impression is that the Management Board’s “approval in Principle” of the reports is final. Appeals appear not to have been an issue during the pilot round but the process should be elaborated in order to avoid future difficulties.
- j. “Provide for consistency of panel decisions.” Much effort has been put into achieving this goal and on the whole it seems to have been successful. In fact, the deliberations and reports seem more consistent than I would have expected in a pilot.
- k. “Be sustainable and affordable for the system as a whole and for the individual colleges in the system.” Sustainability and affordability is a distinguishing characteristic of the audit approach and PQAPA has been implemented in ways consistent with the criterion. I do not believe its costs should be viewed as burdensome—especially since most of the cost at the individual college level should be chalked up to quality improvement and not the audit, per se.

The PQAPA documentation also speaks to the EU’s requirement that the governance of quality agencies should be independent of the institutions being reviewed, a criterion not met formally by OCQAS. My experience with the effects of institutional influence on the USA’s regional accreditors makes me very sensitive to this issue. However, I found no evidence that OCQAS’s ownership by the college system has inhibited its efforts in any way.

Audit Methodology

Like nearly all evaluation methods based on site visits, audit in general and PQAPA in particular involves the following steps: (i) auditee orientation and preparation, (ii) auditor selection and training, (iii) the site visit; (iv) the audit report, and (v) follow-up on the auditors’ recommendations. I’ll review PQAPA’s methodology in each of the categories.

Auditee Orientation and Preparation

This is the area that is most significantly in need of improvement. Both auditee representatives and panel members described the beginnings of the process as confusing. While the broad objectives were clearly stated in the initial documentation, some auditees were uncertain as to whether the process was supposed to review their programs, program review methodology, or QA generally. The guidelines for documentation lacked clarity, which caused institutions to produce too much—thus expending too much effort and inundating the audit panels. It may be that PQAPA’s auditee orientation was more extensive and was apparent to me, but the evidence indicates that improvements are needed.

International best practice puts the same effort into auditee orientation as into auditor selection and training. The process typically starts with a workshop or seminar that addresses the purposes of audit (and the difference between it and assessment), the audit process, and most importantly the design of the self-study and the documentation package. Guidance notes are provided in advance and elaborated in the workshop’s plenary and small-group sessions. While sample reports and packages are difficult to come by in a pilot, they should be provided in future rounds. The

workshops may be held on individual campuses or on a multi-campus basis—e.g., within regions.

Perhaps the key point to stress in the orientation is that preparation for an audit should be viewed as an opportunity for faculty to engage in conversations about teaching and learning quality—conversations that are structured by the audit criteria and forced to closure by deadlines. The goal is not to produce a voluminous package of documentation but instead to reflect upon one’s QA methods and then identify and commit to improvements. These ideas need to be reinforced during the institutions’ preparation cycle as well as being stressed in the orientation. Best practice provides for periodic meetings among those responsible for self-study development and the audit organizers—for example, biweekly or monthly conference calls to identify issues and exemplary approaches and reinforce the audit’s purposes. Bringing together the preparers this way reinforces audit’s improvement agenda and reduces its competitive dimension.

The length of the self-study and the kind of supporting documentation provided should be spelled out clearly in the guidance notes. I believe the self-study’s organization should be spelled out as well, although opinions differ on this. Twenty pages should be more than adequate for the self-study. A list of documents produced and acted upon during the normal course of institutional QA should be included and referred to in the study, but only a few key examples should be included as appendices. The orientation should ask providers to empathize with the auditors’ difficulties in assimilating reams of materials and stress that the goal is to elicit useful conversation during the visit rather than to do a “snow job” on the panel.

Auditor Selection and Training

In contrast to auditee orientation, PQAPA’s auditor training appears consistent with international best practice. The day-long session was appropriately organized, informed by good materials (in particular the *Orientation Manual*), and well led. The use of small groups and a role-play exercise definitely should be retained in future rounds, though the latter might be improved along the lines suggested by some of the participants. The feedback from the session was good.

I suggest that the first-round auditors be invited to participate in future rounds of training. Not only is a “refresher” useful, but the experienced cadres will provide useful information and role modeling for the rookies.

I obtained little information about the auditor selection process, but based on results it appears to have been effective. The practice, adopted by all the teams, of convening regular conference calls during the run-up to the audit is exemplary and should be incorporated in the guidance notes. Such calls identify issues and additional information needs and, importantly, keep the audit at the front of everyone’s mind. So is the practice of dividing the work-load involved in the detailed study of documents.

The Audit Visit

The audit visit generally conformed to best practice: in particular, the drill-down view to “grass-roots” QA activities and the inclusion of students. The ability to get beyond set-piece presentations by senior administrators is crucial, and the teams demonstrated strong awareness of that—for example, by limiting the length of PowerPoint

presentations. The decision by one of the teams to “dis-invite” the president from all but the beginning and ending sessions also was exemplary.

Three-member teams represent the minimum size needed for a competent audit, and while an extra person or two would be helpful the need for economy can understandably take precedence. Splitting the teams to engage more people and issues is essential in a short visit, and this practice should be continued. I understand that individual members provided the chair with notes for inclusion in the report, which should be continued.

The main issue concerns the length of the audit visit. One day strikes me as too short for a campus-level audit. (I’ve run one-day audits in Missouri and Tennessee, but they looked at single departments.) Some of the teams sought to mitigate this difficulty on their own, by agreeing to meet the night before to go over the materials and prepare questions for the visit. Ideally, though, the audit visit would be scheduled for a day and a half rather than one day. This would allow the team to compare notes and identify open issues at a private dinner after the first day. Then they could sleep on their experience before following up on open issues, finalizing their conclusions, and conducting the exit conference the morning of the second day. I know that increasing lengthening the visit always confers benefits and that one must draw the line somewhere. However, adding another half day will produce disproportionate improvements in my opinion.

The Audit Report

The audit reports appear well written and in accordance with the specification that they adhere to the audit principle rather than straying into assessment. PQAPA is to be commended for providing the panels with clear instructions about the form and substance of the report and ensuring the instructions were followed.

The reports are appropriately organized around PQAPA’s five criteria, so the changes in the criteria I’ll suggest below also apply to the audit report. Three additional suggestions follow.

First, PQAPA’s “Guiding Principle (h)” says the audit should “provide for categories of approval”. The reports do this for the individual criteria. Conclusions for the audit as a whole, however, are expressed in broad language: for instance, QA “is well planned and well structured” or the college “is to be commended for its focus on excellence.” Such statements are fine as far as they go, but the overarching evaluation of an organization’s quality assurance and improvement processes can go further. I have found the following five-level categorization scheme, based on the so-called “capability-maturity theory” developed at Carnegie-Mellon university for use in hi-tech industry, to work well in Missouri, Tennessee, and Hong Kong.

- *No effort* means there are no organized quality assurance and improvement (QA&I) processes.
- *Firefighting* means that units respond to problems but mostly with *ad hoc* methods. The QA&I criteria receive little systematic attention.
- *With informal effort*, one sees individual initiatives and experimentation with improvements motivated explicitly by the criteria.

- With *organized effort*, quality process initiatives begin to be planned and tracked, work methods are systematically rooted in the criteria, and the unit has begun to develop performance metrics and norms.
- With *mature effort*, quality processes have been embedded in unit's culture, continuous improvement is a way of life, and organizational learning about QA&I process is fully established.

The categories can be useful in developing summary language for the report even if it should be decided that a formal categorization scheme might expose the exercise to the development of inappropriate league tables.

A second suggestion is to use Australia's three-way scheme in couching the Panel's recommendations.

- *Commendations* call out areas of exemplary practice—areas that are especially good and may be worthy of emulation by other colleges.
- *Affirmations* call out areas where the college itself has found a weakness and intends to correct it. In effect, the Panel is affirming the college's judgment.
- *Recommendations* call out areas needing improvement that the Panel has discovered on its own.

The *Guidance Notes* should make clear that affirmations are better than recommendations. In addition to singling out areas of best practice, the scheme provides an incentive for respondents to identify and disclose weaknesses themselves rather than hoping the Panel won't find them.

Finally, the *Notes* state that "it is assumed that all colleges will extend their tradition of openness and transparency" to their QA&I processes. I agree this is essential. It also follows from the statement that the audit reports and institutions' responses should be made public, but I'm not sure whether there is a firm policy to disclose. The need for disclosure stems from more than public accountability. Being an effective stimulus for change requires that the report become a touchstone on campus, which in turn requires wide distribution and much discussion. Such discussion is sure to become public knowledge and, depending on the content, the target of media interest. Therefore, it's wise to release the report and resulting plans for improvement at the outset.

Follow-Up

International best practice provides for systematic follow-up on the audit recommendations in a more timely way than during the next audit cycle. I believe OCQAS's follow-up processes are still in a formative stage and that it will be important to codify the approach before the next round gets under way.

One cost-effective method of follow-up is to get a written report from the institution a year to eighteen months after the audit. The report should describe what the college did to remedy the defect, when, what the results were, and what additional actions are planned. The reports would be reviewed by the OCQAS Management Board, who would decide whether additional action is needed. (Such action might involve request for additional information, a visit from OCQAS staff, or perhaps even a formal communication to the president.) The follow-up report and any ensuing documentation would be provided to the next cycle's audit panel.

The identification and sharing of exemplary practices is an important byproduct of audit. Therefore, follow-up sessions to share exemplary practices like the one I attended at George Brown College should be continued.

Quality Assurance and Improvement Criteria

PQAPA is built around five criteria that “define institutional policies and practices that a college will have developed and implemented to ensure the quality of their programs of instruction.” Such criteria provide guidance to the institutions in developing their QA&I programs and in preparing for the audit. They also provide the auditors with standards against which to judge what they see. In other words, they lie at the core of the audit process.

The criteria are well conceived and by and large they have served as intended. Their careful statement of requirements followed by explanatory notes is worthy of commendation. However, reading the reports, listening to testimony from the auditors and participating institutions, and studying the criteria in light of my own knowledge about QA&I processes lead me to suggest some changes. In particular, the criteria as stated intermingle certain key concepts, contain redundancies and seem to project a slightly more “procedural” flavor than I believe was intended.

It seems to me that the list should begin by asking whether a clear statement of learning objectives exists for the program. Subsidiary questions include the kinds of evidence that were taken into account in developing the objectives—e.g., evidence of student and employer need as well as discipline-oriented content knowledge—and whether they are updated regularly. Further tests include whether the objectives are “operationally meaningful” in the sense of providing guidance for curriculum development and student learning assessment, and whether program faculty have internalized and use them in their day-to-day work. (An important test of meaningfulness is whether achievement of the objectives can be measured—objectives for which achievement cannot be measured, at least in principle, are not very useful.) These are auditable standards that do not run to the question of whether the objectives are “appropriate” in the eyes of the auditors.

The current *Criterion 1*, dealing with applicant and student requirements, will become more meaningful and auditable once the groundwork in terms of program objectives has been laid. The first question is whether the requirements flow coherently and reasonably from the objectives. This can be followed by the questions about clarity, etc., that comprise the current statement.

This is a good place to note a subtle difference between what’s proposed and the existing criteria. The criteria as stated sometimes seem to emphasize structures and documents—in other words, formalities—more than use of evidence, seriousness of attention to QA&I, and other aspects of behavior. This is understandable in light of the need to avoid judging “quality” as opposed to “quality assurance,” and also because of audit’s roots in the early 1990s in the UK, but it misses some important matters. Fortunately, both the institutions and the audit teams seemed to go beyond the black letter of the criteria in doing their work.

The current *Criterion 2* speaks to how the program objectives are mapped into the curriculum and credentialing. The idea of mapping might be made more explicit and extended to include the use of evidence—e.g., why were choices made as they were and how do the respondents judge their success in adhering to the objectives. (This

provides another example of how the audit, by going beyond formalities, can spur dialog and improvement.) Finally, *Criterion 1.2* seems to overlap with *Criterion 2*—which is one source of the redundancy mentioned above.

Criterion 3 could fruitfully emphasize the development and continuous improvement of teaching and learning methods to a greater extent than at present. The present *Criterion 3.1* covers this but at a very high level—furthermore, the explanation simply restates the requirement. An elaborated requirement might stress the need for regular experimentation with new teaching methods with the results being measured as systematically as possible. It might also ask about how efficiency in the use of resources is considered in the design of teaching methods.

Criterion 3.4 appears redundant with *Criterion 5*. It's true that good teaching includes giving feedback to students but it's not clear that is what's meant. *Criterion 5* aims at getting feedback to faculty and the institution. It should call for a continuous improvement process, which could include some of what's in *3.4*, rather than something walled off in periodic program review “episodes”. *Criterion 5* might also be revised to call for use of evidence in regard to *Criteria 1-4*. This introduces some redundancy, but in this case I think it's worthwhile to reemphasize the all-encompassing and seamless nature of assessment.

I recognize that PQAPA has a considerable investment in the criteria as they stand, and thus I've not tried to rewrite them. My hope is that these suggestions will trigger dialog that leads to whatever improvements eventually are deemed appropriate.

In closing, let me say again how impressed I am with the PQAPA pilot. I've enjoyed learning about it and meeting the participants, and I wish everyone well.

Respectfully submitted,

William F. Massy