Measuring the performance of professional regulatory bodies: A review of current models and practices

Claude Balthazard, Ph.D., C.Psych., CHRL
VP Regulatory Affairs and Registrar
Human Resources Professionals Association

January 11, 2017

This report gives an overview of current approaches and thinking about the measurement of performance of professional regulatory bodies.

Measuring the performance of professional regulatory bodies is not an easy task. There are also different definitions of what performance means. The bottom line for professional regulatory bodies is the extent to which potential harms to the public and other users of the professional service have been reduced, suppressed, mitigated or eliminated. This harm-reduction is the ultimate outcome for all professional regulatory bodies. For many reasons, measuring the degree to which harms have been reduced, suppressed, mitigated or eliminated is not an easy thing to do and is rarely seen in practice. In most cases, ‘measuring performance’ means measuring not only results but other aspects which lead to results.

Professional regulatory bodies are different

Performance measurement is one of those endeavours that differs significantly depending on the type of organization. Although generalizations are sometimes possible, this must be done carefully. For instance, professional self-regulating bodies are different from other kinds of regulators, they are different from government agencies, they are different from other kinds of not-for-profit corporations, and they are different from for-profit corporations. It is also the case that professional regulatory bodies for voluntary professions are somewhat different than professional regulatory bodies for licensed professions. The implication is that performance assessment methodologies developed for for-profit corporations, non-regulatory not-for-profit organizations, and regulators other than professional regulators must be adapted or translated if they are to be applied to self-regulating professional bodies.
Specific measurement challenges in measuring the performance of professional regulatory bodies

1. The ultimate objective for any professional regulatory body is the promotion and protection of the public interest, but each professional regulatory body is left to define what this means in their context and for their profession.

2. Professional regulators are in the harm prevention business—in other words, performance means that harms do not happen. As pointed out by Sparrow, it is difficult to count what has not happened.

3. Professional regulation is inherently a complex endeavour. Regulated professionals are not the ‘customers’ for the professional regulatory body’s product or service—the public is. Professional regulatory bodies are supported by the profession but are there to serve the public interest. This complicates matters.

Different approaches to the measurement of the performance of professional regulatory bodies

No single approach has won the day—no one has ‘cracked the code’ of performance measurement for professional regulatory bodies. The trend has been for professional regulatory bodies to adapt methodologies first designed for for-profit corporations or for non-regulatory not-for-profit corporations. For instance, the logic model was first developed for social benefit not-for-profit organizations and the balanced scorecard was developed for for-profit businesses. It somewhat late in the game that approaches designed specifically for professional regulatory bodies have emerged. For instance, the Professional Standards Authority’s Standards of Good Regulation first appeared in 2010. Malcolm Sparrow (2008) makes the argument that harm-reduction—which is the business that regulators including professional regulatory bodies are in—is fundamentally different from other endeavours and that, consequently, a different approach to performance evaluation is required for regulatory bodies.

Measurement for what purpose?

A number of sources have pointed out that different measurement approaches are required for different purposes. The implication here is that the purposes to which the measurement system will serve need to be clear at the outset.

The College of Physiotherapists of Ontario referred to five types of balanced scorecards based on five different purposes.

1. Valuation – most often shareholder facing
2. Navigation – monthly, forward facing, internal focus
3. Compensation – outcome focused
4. Benchmarking – comparing against performance of others, external focus
5. Evaluation – annual, reporting upstream

The College of Physiotherapists of Ontario indicated that their balanced scorecard was of the navigation type.

Coglianese (2015) noted:

“*The proper design of any regulatory performance measurement system will depend ultimately on its purpose. Measurement is a good tool for learning, and learning should*
be a core part of the excellent regulator’s culture. But the point is not merely learning for learning’s sake. Regulatory leaders need to be strategic about measurement. They need to determine clearly the purposes – internal and external – for measurement, and then plan a system that meets that purpose. There exists no one-size-fits-all “best” measurement system, and no single measure can capture all that a regulator must accomplish to become excellent nor all that a regulator might seek to learn with measurement.”

And:

“The broadest approach to measurement might be an overall “excellence assessment.” Such an assessment of excellence is different from the strategic use of measurement any excellent regulator puts in place in order to run its operations well. Measurement of excellence, in other words, is different from measurement for excellence, and it may call for a different approach to measurement.”

There are a number of factors which go into which approach is ‘best’ including the culture of the professional regulatory bodies and the purpose of the measurement program. Below is a list of different purposes for a professional regulatory body to develop and maintain a measurement program:

- To assess the performance of the regulatory body CEO and/or other regulatory body executives (‘pay for performance’)
- A mechanism for the Board to demonstrate that it is discharging itself of its statutory accountabilities (‘oversight’)
- To satisfy the government that the professional regulatory body is fulfilling its statutory mandate to promote and protect the public interest
- To be able to demonstrate to the public and other stakeholders that the professional regulatory body is effective in promoting and protecting the public interest
- To assist the professional regulatory body in the pursuit of regulatory excellence
- To assist in elaborating the professional regulatory body’s strategy
- To maximize the longer-term impact of the professional regulatory body by providing information that will shape future priorities and investments
- As a tool to help the professional regulatory body to focus on the right things
- To provide a means of tracking progress against operational targets
- To determine which initiatives are effective and which are not
- To provide a tool to compare the organization’s performance against that of other professional regulatory bodies (‘benchmarking’)

Six approaches to performance measurement for professional regulatory bodies

Six approaches to performance measurement for professional regulatory bodies are reviewed:

1. Activity and output based measurement
2. The logic model and theories of change
3. The balanced scorecard
1. Activity or output based measurement

The first approach is the common activity or output based measurement which is the typical approach to the measurement of the performance of professional regulatory bodies. The focus is on activities or outputs. The focus is on ‘how many,’ ‘how quickly,’ or ‘how many per given amount of dedicated resources’ (efficiency). These activity or output measures are easy to measure but do not give an indication of the effectiveness or impact of these activities.

Outputs should not be confused with outcomes. Outputs refer to counting activity in terms of applications processes, exams administered, or complaints disposed of. Outcomes refer to the impact that the activities have had.

For example, most measures to be found in HRPA’s Registrar’s Report are for the most part of the activity or output type. Typical measures are:

- Number of registration applications processed
- Number of exams administered
- Number of designations granted
- Number of complaints received
- Number of complaints disposed of
- The average number of days it has taken to dispose of a complaint
- The number of referrals to discipline

Interestingly, of the measures above there are some that are not necessarily activity or output metrics. In a sense, there is a difference between ‘active’ activities or outputs and ‘reactive’ activity or outputs. For instance, the number of complaints received does say something about how busy staff and volunteers will be in processing these complaints, but filing a complaint is something that the public does (although in most professional regulatory bodies there is a mechanism for ‘internal’ complaints). With the idea that complaints processed will be the same as the number of complaints received (other than lag or backlog), the number of complaints processed is not something that the professional regulatory body can control directly. As such the number of complaints filed could be seen as an outcome measure. Similarly, with registration—the number of applications for registration processed is an activity or output measure, but the number of applications for registration received could be considered an outcome measure.
Compliance also illustrates how activity and outcomes can be linked as chains.

Although ‘objective,’ some measurements are difficult to interpret. For instance, what meaning should be given to a low number of complaints?

- Is it that the professional regulator is doing such a good job of managing the risks posed by the practice of the profession that the incidence of harms is very low?
- Is it that the actual incidence of professional misconduct in the profession is very low and has nothing to do with the actions of the professional regulatory body? (One could then question the need for professional regulation and a professional regulatory body in the first place.)
- Is it that the incidence of professional misconduct is actually significantly higher but that such professional misconduct is under-reported for any of a number of reasons such as:
  - The fear of retaliation on the part of the professional who would be the subject of a complaint,
  - The public doesn’t know that the regulated professionals are in fact regulated and are accountable for their behaviour to a professional regulatory body,
  - The public doesn’t know how to file a complaint against a member of the profession,
  - The public has little or no confidence that the regulator will act in the public interest but believe that the professional regulatory body will act to ‘protect its own.’

Similarly, a high level of complaints is equally open to alternative interpretations. Does it indicate that the professional regulatory body is vigilant in prosecuting professional misconduct on the part of their
registrants or does it indicate a failure in programs intended to prevent professional misconduct in the first place? Determining what the number of complaints should be would require research.

The main criticism of activity or output measures is that they say little or nothing about how well the professional regulatory body is delivering on its mandate to promote and protect the public interest—in other words, business is not effectiveness. This is the difference between activity or output and outcome. For instance, the professional regulatory body may engage in a number of activities designed to facilitate the filing of appropriate complaints on the part of the public, but such activity may be ineffective.

The second criticism of activity or outcome measures is that they tend to favour amount rather than quality. For instance, in regards to investigations, counting the number of investigations or the amount of time it takes to complete investigations may lead to less concern for the thoroughness of the investigations. In other words, quality, which is usually more difficult to measure than quantity, but which may have more impact on outcomes than quantity will get ‘short shrift.’

The problems with activity or output measures are:

1. Being ‘busy’ is not the same thing as being effective.
2. What is easily quantified may not be what is most important

Alternatives to activity and/or output measurement

A number of approaches have been discussed or used to get beyond simply measuring activity or outputs.

2. The logic model and theories of change

A logic model is a tool used by funders, managers, and evaluators of programs to evaluate the effectiveness of a program. Logic models are usually a graphical depiction of the logical relationships between the resources, activities, outputs and outcomes of a program. As far as this author is aware, no professional regulatory body specifically references the logic model—however, the model is so basic and so intuitive that its influence can be seen in many places.

The early work on the logic model was done by Carol Weiss in the 1970s. Since then many refinements and variations have been added to the basic concept. Many versions of logic models set out a series of outcomes/impacts, explaining in more detail the logic of how an intervention contributes to intended or observed results. This will often include distinguishing between short-term, medium-term and long-term results, and between direct and indirect results. Logic models are used both for program development and for program evaluation.

Here is an often cited version from the Kellogg Foundation Logic Model Development Guide (2004).
The logic model approach is also the approach used in the federal government to evaluate programs.

Of course, actual logic models are somewhat more complicated. Below, is the sample logic model included in the Government of Canada Guide.
Logic models and theories of change

Another methodology which is linked to the logic model approach is the theory of change. The terms logic model and theory of change are often used interchangeably although they are distinct but related methodologies. The table below was created based on a presentation by Clark & Anderson (2004).

The Theory of Change methodology comprises the following steps:

1. Identifying long-term goals and the assumptions behind them
2. Backwards mapping from the long-term goal by working out the preconditions or requirements necessary to achieve that goal—and explaining why.
3. Stating the assumptions about what exists in the system without which the theory won’t work, and articulating rationales for why outcomes are necessary preconditions to other outcomes.
4. Weighing and choosing the most strategic interventions to bring about the desired change.
5. Developing indicators to measure progress towards desired outcomes and assess the performance of the initiative.
6. Reviewing the theory to ensure that it is: (1) plausible, (2) “doable” (or feasible), and (3) testable.
7. Writing a narrative to explain the summary logic of your initiative.

The Theory of Change is often presented graphically as an outcomes framework, an example of which is given on the next page.
Logic models | Theories of Change
---|---
**What they are** | Graphically illustrates program components, and creating one helps stakeholders clearly identify outcomes, inputs and activities | Links outcomes and activities to explain HOW and WHY the desired change is expected to come about
**Starting point** | Usually, starts with a program and illustrate its components | May start with a program, but are best when starting with a goal, before deciding what programmatic approaches are needed
**What they require** | Requires identifying program components, so you can see at a glance if outcomes are out of sync with inputs and activities, but they don’t show WHY activities are expected to produce outcomes | Also, requires justifications at each step – you have to articulate the hypothesis about why something will cause something else (it’s a causal model, remember!)
**Indicators** | Doesn’t always identify indicators (evidence to measure whether outcomes are met or not) | Requires identifying indicators

One way to think of the relation of logic models to theories of change is that the theories of change are a ‘deep dive’ into the causal linkages that underlie a logic model. The two methodologies can work together—after having worked out a theory of change, a logic model can act as a summary of this underlying model.

**Summary for logic models and theories of change**

As noted earlier, logic models and theories of change seem rarely used by professional regulatory bodies. There are likely many reasons why this is the case.

- Much of the development of logic models and theories of change was a result of funders increasingly expecting agencies to demonstrate the impact of the programs they were seeking funding for. Professional regulatory bodies are not subject to the same pressures. For instance, if the continued existence of professional regulatory bodies was tied to the demonstration of impact, the situation would likely be quite different.
- Professional regulatory bodies do not start with a blank slate. Enabling legislation is often quite specific about what professional regulatory bodies must do. Although professional regulatory bodies are free to go beyond what they are strictly mandated to do (subject to the limits of their stator powers), there is already a theory of change in place.

The logic model approach is similar to the balanced scorecard which will be taken up next.

The logic model does not solve all problems:

- As pointed out by Sparrow (2008), mapping out the last step—from outputs to outcomes or impacts—is particularly difficult for regulatory bodies.
Example of an Outcomes Framework
• Logic models are best suited to programs, assessing the performance of a regulatory body would require multiple logic models or complex logic models.

3. The balanced scorecard

Another popular causal model is the balanced scorecard (or some derivation thereof). The original balanced scorecard was developed for for-profit businesses, but adaptations were made for not-for-profit organizations and government.

The basic idea was to move away from measuring solely the bottom line and to move back into the causal chain of what delivers bottom-line performance. The balanced scorecard was originated by Robert Kaplan and David Norton in the early 1990’s as a performance measurement framework that added strategic non-financial performance measures to traditional financial metrics to give managers and executives a more 'balanced' view of organizational performance.

For professional regulatory bodies, there are a few important differences, however. The bottom line is not financial but the promotion and protection of the public interest. With self-regulation, the funds to support the regulatory enterprise comes from the professionals under regulation but the ‘customer’ is the public (or the Legislature on behalf of the public).

An important aspect of the balanced scorecard approach is the development of key performance indicators for the various components of the causal chain. The balanced scorecard approach puts a heavy emphasis on quantifying each component of the causal chain but does not propose any new approach to the quantification of these components. As a result, balanced scorecards tend to make use of measures which already exist in the organization. Often the balanced scorecard becomes just another way of organizing or presenting the same measures that were in place before the introduction of a balanced scorecard.

Examples of the use of the balanced scorecard by professional regulatory bodies:

• College of Physicians and Surgeons of Ontario, Strategic and operational dashboard, 2016
• College of Medical Radiation Technologists of Ontario, Balanced scorecard and dashboard, 2015

The best-known example of the use of a balanced scorecard is the one used by the College of Physiotherapists of Ontario.
One criticism of the balanced scorecard approach—at least as it is commonly applied to professional regulatory bodies—is that the key performance indicators for outcomes are not what most would think of outcome measures. Consider for example the key performance indicators for the public protection quadrant in the College of Physiotherapists balanced scorecard:

- Number of notices issued
- Number of appeals heard
- Number of mediated resolutions to complaints
- Number of complaints disposed of
- Number of discipline decisions issued
- Number of quality and safety referrals
- Number of interprofessional collaboration initiatives
- Patient surveys results
- Communications campaign initiatives
- Public reporting

For the most part, these are measures of activities or outputs related to the protection of the public but do not measure public protection outcomes (i.e., harms reduced, suppressed, mitigated or eliminated). What seems to happen in practice is that balanced scorecards become ways of organizing activity measures.
The balanced scorecard and logic models

The balanced scorecard is similar in many respects to the logic model. Interestingly enough, these seem to be two different worlds—those working within the logic model framework rarely make reference to the balanced scorecard and those working within the balanced scorecard framework rarely make reference to the logic model.

There are some sources which describe how the balanced scorecard and the logic model relate to each other. For instance, in Creating Program Logic Models: A Toolkit for State Flex Programs, Gale, Loux, & Coburn (2006) describe how the balanced scorecard and the logic model can complement each other. According to Gale, Loux, & Coburn (2006), one weakness of the balanced scorecard is that it is based on an already developed strategic plan which assumes that the organization’s strategic plan and related strategies and activities are clearly aligned with the achievement of its stated goals and objectives. These authors note that “this is not always the case, particularly when an organization has failed to make explicit the underlying program theory upon which its strategic plan and activities are based.” They go on to state:

“The creation of an organization- or program-specific logic model can help to address this problem and can serve as an important developmental step in the implementation of a balanced scorecard by clearly laying out the organization’s or program’s goals; mapping the theory, assumptions, activities and strategies necessary to achieve those goals; and identifying appropriate outcomes and indicators to map progress towards the achievement of those goals.”

To summarize, the balanced scorecard does bring the focus to the causal antecedents of ultimate outcomes but it does not of itself solve some of the difficult aspects of measuring the performance of professional regulatory bodies.

4. Models based on best practices

Practice-based measurement is based on the idea that doing the right things in the right way will lead to results. This approach focuses much less on counting activities or outputs and more on what is being done and how well it is done. This approach is flexible and adaptable. However, the validity of the practices is key (how the practices were identified). Practices are identified based on a model of what is effective. This model may be quite plausible but it will rarely be based on empirical data. The ‘organizing framework’ for the practices will also vary. For instance, the Professional Standards Authority in the UK has adopted a model organized by regulatory function whereas Coglianese’s model is organized more on a process model. There is the issue of level, does the practice reflect a minimally acceptable level of performance or an excellent level of performance. The strength of the linkage between a given practice and outcomes can be difficult to establish. As the impact of any practice on outcomes will likely depend on the presence of other practices, the impact of practices upon outcomes...
is likely to be wildly non-additive. The evaluation as to whether a practice is in place or the extent to which a practice is in place is usually based on a rating process which can be more or less subjective.

Examples:
- Royal College of Dental Surgeons of Ontario
- College of Registered Nurses of British Columbia
- The College of Early Childhood Educators of Ontario
- Real Estate Council of Alberta
- HRPA Gold Standard Audit (2013)
- The Office of the Fairness Commissioner of Ontario

There are different variations of this approach. In some approaches, the practice defines a minimally acceptable level of performance. Professional regulatory bodies are scored on a meet/does not meet basis for each practice. This is the approach taken by the Professional Standards Authority in the UK. The Professional Standards Authority in the UK is a meta-regulator (it regulates regulators).

The approach of the Professional Standards Authority in the UK is organized around four regulatory functions: guidance and standards, education and training, registration, and fitness to practice. In its most recent formulation (January 2016), the Professional Standards Authority’s Standards of Good Regulation comprise twenty-four practices. The Professional Standards Authority’s approach has been quite popular. The Professional Standards Authority used their framework to assess the Royal College of Dental Surgeons of Ontario (2012) and the College of Registered Nurses of British Columbia (2016). Also, the Professional Standards Authority framework was used as the basis of a self-assessment by the College of Early Childhood Educators of Ontario and for a third-party assessment by the Real Estate Council of Alberta.

HRPA’s Gold Standard Audit (2013) defined two performance anchors for each practice—one at the minimally acceptable level of performance and the other at the excellent level of performance. With ‘in-between’ scores allowed, this gives a 5-point measurement scale for each practice. Although developed independently of the Professional Standards Authority framework, both frameworks cover much of the same domains; however, the Gold Standard Audit covers more of the cultural aspects which are not tied to any specific regulatory function. In its most recent formulation, HRPA’s Gold Standard Audit defined 95 practices. One difference between the Professional Standards Authority’s Standards of good regulation and HRPA’s Gold Standard Audit is that the latter defines many more practices related to organizational culture and governance.

The Office of the Fairness Commissioner of Ontario (OFC) took a similar approach to assessing compliance with the Fair Registration Practices Code. It should be noted that the OFC’s concern was narrower. The OFC did not seek to assess performance as a professional regulatory body but to assess compliance with the Fair Registration Practices Code. This mandate also provided the organizing framework for the assessment. The OFC assessment was organized according to the duties spelled out in legislation.

The OFC has three anchors or ratings for each practice:
- Demonstrated — all required elements of the practice are present or addressed
- Partially Demonstrated — some but not all required elements are present or addressed
• Not Demonstrated — none of the required elements are present or addressed

Plus a fourth rating of:
• Not Applicable — this practice does not apply to the regulatory body’s registration practices

Practice-based models have some similarity to activity-based measurement but instead of focusing on ‘how much’ or ‘how quickly,’ practices-based measurement focuses on the degree to which a practice is in place in the professional regulatory body. Doing the right thing in the right way is more important than doing a lot of the wrong things.

Summary of practice-based approaches

Best practices approaches do not measure outcomes. The measurement of practices is predicated on the idea that certain practices are related to desired outcomes. The validity of the practices is always an issue. What evidence is there that the incorporation of a given practice will necessarily lead to the desired outcome? It is rare that the link between a practice and performance outcome is based on empirical evidence.

Another issue is that the measurement of practices is not about counting things but of rating as to whether actual practices meet a given standard. As such, practice-based measurement always involves judgments made by judges. To some, this introduces an undesirable level of subjectivity in the ratings.

5. Coglianese 2012 and 2015

Recently, Cary Coglianese has made two contributions to the measurement of the performance of regulatory bodies. The first is a paper which was commissioned by the Organisation for Economic Co-operation and Development (OECD) which was published in 2012. The second was a project commissioned by the Alberta Energy Regulator (AER) to define the characteristics of ‘best in class’ regulatory agencies in 2015.

At this point in time, there are no published accounts of Coglianese’s models being used to assess the performance of a professional regulatory body; nonetheless, Coglianese’s work deserves mention because it is focused specifically on the measurement of the performance of regulators and the amount of research that went into the development of the regulatory excellence model.

Coglianese’s 2012 effort was in a program evaluation approach to the work of regulators:

“This paper develops a framework for systematically evaluating the performance of regulations and regulatory policies. Offering an accessible account of the fundamentals of evaluation, the paper explains the need for indicators to measure relevant outcomes of concern and research designs to support inferences about the extent to which a regulation or regulatory policy under evaluation has actually caused any change in the measured outcomes. Indicators depend on the specific problems of concern to policymakers as well as on data availability, but the best indicators will generally be those that measure the ultimate problem the regulation or policy was intended to solve. In addition, research designs should seek to emulate the structure of laboratory experiments in order to permit valid causal inferences about the impacts of a regulation
A feature of Coglianese’s model is the role of behaviour change in the entities under regulation. This idea applies well to professional regulation where the idea is to ‘govern and regulate the practice of members of the Association and firms.’ The idea is that professional regulatory bodies are about behaviour modification—modifying the behaviours of professionals that have the potential to create harm for clients and users of the professional service. The evaluation of the effectiveness of a professional regulatory body would require the professional regulatory body to have a list of targeted behaviours.

Coglianese’s model differentiates between different levels of outcomes: immediate outcomes, intermediate outcomes, and ultimate outcomes. The immediate outcomes are defined as behavioural changes in the regulated entity.

For instance, in Human Resources, consider the harms that may come from the misuse of employment tests. The Ultimate Outcome of Concern would be protection of the public; the Intermediate Outcomes of Concern would be outcomes such as the denial of career opportunities to those who deserve them, loss of productivity, and dysfunctional workplaces; the behavioural change would be the competent and effective use of employment tests; the Regulation of Interest may be a combination of amendments to the prescribed curriculum, changes to the certification processes, the issuance of practice guidelines in regards to employment testing, and perhaps the addition of mandatory content to the Continued Professional Development requirement.

In 2015, Coglianese was commissioned to look into what makes best-in-class regulators. Based on extensive research, he developed a regulatory excellence model based on three core attributes.
In the context of this work, Coglianese proposed the simplified model below. This later model retains the ‘activities-behaviours-outcomes’ sequence, but it adds an organizational component at the front end (like the balanced scorecard), and a perceptual outcomes loop. The latter reflects the fact that public confidence in the willingness and ability of the regulator to regulate in the public interest is an important factor that reinforces the capacity of the regulator to influence behaviour.

**MODEL OF REGULATORY ORGANIZATION, ACTION, AND PERFORMANCE**

- **ORGANIZATION**
  - The Regulator
  - People (Internal Management)

- **ACTIONS**
  - Decisions → Actions
  - Public (External Engagement)
  - Priority-Setting
  - Problem-solving

- **BEHAVIOR**
  - The Regulated
  - Intermediate → Ultimate

- **Substantive OUTCOMES**
  - E.g., Approval, Trust, Legitimacy

- **Perceptual OUTCOMES**


In regards to measurement, Coglianese is somewhat eclectic—“excellent regulators need a range of measures, narrow and broad, on the regulator as well as on the world.” Coglianese thinks in terms of a portfolio of performance measures. This portfolio of performance measure would include some practice-based measures and some outcome-based measures.

“To use measurement for the purpose of achieving excellence, a regulator should include in its measurement portfolio some evaluation research that aims to trace out causal connections between the regulator’s actions and outcomes in the world.”

Notwithstanding, for the most part, Coglianese uses a practice-based approach to assess regulatory performance. Coglianese developed an “excellent” checklist for regulators’ which is comprised of twenty-four practices. These 24 practices are organized under four headings: (1) internal management (e.g., mission clarity, resources, autonomy, human capital, and culture), (2) priority-setting/decision-making (e.g., scientific and economic analysis and how it informs decisions), (3) problem-solving (e.g., regulatory instrument design, inspection and enforcement strategies), and (4) external engagement (e.g., transparency, public engagement).
It is interesting to note how Coglianese’s twenty-four practices are quite different from the Professional Standards Authority’s twenty-four practices. Can the 24 practices of the Professional Standards Authority be reconciled with the 24 practices of Coglianese? The frame of reference for Coglianese’s work was regulation in general and not professional regulation specifically. The Professional Standards Authority framework was developed specifically to assess the performance of professional regulatory bodies. Nonetheless, these two approaches could be seen as two different ways of carving up the same pie. Some professional regulatory bodies might find one or the other way of carving up the same pie more useful. There is no reason why both could not be used. In fact, HRPA’s Gold Standard Audit approach, with 95 practices, included practices relating to specific regulatory functions (as per the Professional Standards Authority) and practices relating to internal processes (as per Coglianese).

6. **Sparrow’s harm-prevention model**

Finally, a somewhat different kind of causal model. This model focuses not on the regulatory body per se but the harm and its ‘chronology.’ The causal chain is not organized around the regulator and its activities but around the harm and its chronology. Sparrow’s approach is that the role of the regulator is to sabotage harms before they happen by intervening early in the chronology of a harm.

Sparrow’s approach challenges some basic tenets of how regulation should be done. In a nutshell, Sparrow’s approach is as follows:

- Regulators should be able to give an account of their performance in terms of harms reduced, suppressed, mitigated or eliminated.
- ‘To give an account of one’s performance’ as regulators means being able to tell a credible story as to how the actions of the regulator have reduced, suppressed, mitigated or eliminated harms.
- A performance story has five components:
  1. The data and analysis which led the department to focus on this particular phenomenon,
  2. Metrics the department put in place and benchmarked up front, so they could tell if they made any headway in reducing the problem,
  3. A description of the intervention implemented,
  4. The impact, measured in terms of changes in the designated project, and
5. An account of any resulting decision regarding project closure, coupled with any longer-term monitoring or maintenance plan put in place to prevent the problem, once suppressed, from re-emerging.

- “If the bulk of the work done, however, is organized around functions (such as enforcement), programs (e.g. partnership programs) or processes -- rather than around identified concentrations of harm -- then a compelling account of harms controlled will remain frustratingly elusive.” The work of regulators should be organized around identified concentrations of harm.

- Several thorny issues relating to these remain largely unresolved:
  (a) Is it possible to prove causality?
  (b) Is it possible to assign credit for harm-reduction to particular functions or programs?
  (c) Is it possible to measure prevention? Can anyone measure accidents that didn't happen, or count the criminals averted from their purpose or deterred?

Sparrow’s approach is challenging because it is not simply another approach to the measurement of performance as a regulatory body but requires a reorganization of the regulatory body in terms of identified harms. Sparrow’s approach puts an emphasis on the identification of potential harms.

One issue with these narrow approaches is that professional regulatory bodies are expected to protect the public from a whole host of potential harms, some of which may not yet have been identified. Using Sparrow’s approach, assessing the performance of a professional regulatory body would require summing the performance over numerous harm clusters.

**The strategic and the tactical**

It would be possible to reconcile Sparrow’s harm-avoidance model with functional models or process models by thinking about the difference between strategic and tactical approaches. Now there are many different ways of defining ‘strategic’ and ‘tactical,’ the distinction here is one that is sometimes used in the game of chess. Strategic moves are those that are made without any specific sequence of moves in mind. For instance, ‘effective deployment’ and ‘controlling key diagonals’ are strategic thinking. At the time the move is made, it is not known how the resulting position will be useful at some later stage. A tactic, on the other hand, is a specific sequence of moves with a specific outcome in mind. In this approach, strategy works at the ‘meta’ level.

This distinction can be applied to professional regulation. Sparrow’s ‘chronologies of harms’ are specific sequences of events which can result in harm to the public. The sabotage of harm can be understood in terms of tactics. The counterpart of strategic thinking would be to ‘develop the harm identification capabilities’ or ‘enhance the professionalism of registered professionals.’ The idea is that these latter activities create conditions whereby the professional regulatory body can develop and implement specific harm avoidance initiatives.

**Another perspective**

It should be noted that professional regulatory bodies do not have a complete choice as to how to fulfil their objects. Indeed, there is a causal model of sorts already embedded in the enabling legislation.

Consider, for example, the *Registered Human Resources Professionals Act, 2013*. Not only did the *Registered Human Resources Professionals Act, 2013*, provide HRPA with a mission—to ‘promote and
protect the public interest by governing and regulating the practice of its members and firms”—it also was quite specific in the activities that must be included in delivering on this mandate.

Specifically, HRPA must:

1. establish, maintain, develop and enforce standards of qualification,
2. establish, maintain, develop and enforce standards of practice,
3. establish, maintain, develop and enforce standards of professional ethics,
4. establish, maintain, develop and enforce standards of knowledge, skill and proficiency, and
5. regulate the practice, competence and professional conduct of members of the Association and firms;

For better or for worse, the Registered Human Resources Professionals Act, 2013, embeds a particular model of regulatory performance. One could define performance as a professional regulatory body doing exactly what our enabling legislation requires us to do. It is not enough for HRPA to demonstrate that it had promoted and protected the public interest by effectively governing and regulating the practice of its members and firms, it must be able to show that it did what its enabling legislation required it to do. On the other hand, our enabling statute does include an object— to address any other matter that relates to the regulation of its members that the Board considers appropriate—which enables going beyond establishing, maintaining, developing and enforcing professional standards.

It is not surprising, then, that many regulators focus on activities since these activities are specifically referred to in their enabling legislation.

Complications specific to voluntary professions
An important aspect of self-regulation is that the direct beneficiaries of professional regulation are the public but the costs of regulation are borne by the regulated professionals. In business, the model is simpler—businesses are successful to the extent that they deliver value to their customers. In professional regulation, whether direct regulation by government or by self-regulation, the ‘customer’ is not the regulated professional but the public (or the Minister or the Legislature on behalf of the public). In fulfilling their statutory mandates, professional regulatory bodies must do some things that may not be popular among amongst the professionals they regulate. As the then Institute of Chartered Accountants of Ontario put it in a delegation to the Standing Committee on Social Policy “We are a regulatory body. Frankly, if our members are happy with what we’re doing, we’re not doing our job.” Professional regulatory bodies should not think of the professionals they regulate as ‘customers’—at least not in any usual sense of the term ‘customer.’

This complicates the professional regulation ‘business model’ for self-regulated professions. Professional regulatory bodies need the support of the professionals they regulate but do not exist to serve the professionals they regulate. This is true of both voluntary and licensed professions, although some of the issues may be greater for voluntary professions.

The point is that this complexity in the voluntary self-regulation business model must appear in the performance causal model.

**Overall summary and recommendations**

Performance means different things to different people. To some performance means outcomes and impact, to others it also refers to all aspects leading up to and including outcomes and impact.

As far as this author is able to tell, no professional regulatory body in Ontario has assessed its effectiveness in terms of overall impact. The measurement of impact or ultimate outcomes is just not seen with professional regulatory bodies. In part, this is no doubt a reflection of the methodological difficulties in measuring outcomes in professional regulation. In most cases, performance measurement is limited to activity or output measures, and in some cases in terms of practices.

Amongst professional regulatory bodies, the current state of the art seems to be practices-based assessments of which the Professional Standards Authority’s Standards of Good Regulation seems to be the most well-known. Authors such as Malcolm Sparrow will fill the room at presentations and workshops but there is no evidence that his ideas are changing the way professional regulatory bodies assess their performance.

There are some balanced scorecards although it is not clear that these are entirely faithful implementations of the balanced scorecard.

The level of sophistication among professional regulatory bodies in regards to performance measurement would appear to be rather basic. There is no published or readily accessible evidence that a professional regulatory body has worked through a logic model or worked out a theory of change.

The following conclusions can be drawn:

- There is no single approach to the measurement of performance for professional regulatory bodies which has won the day.
• The ‘best’ approach will depend on a number of factors including the purpose of performance measurement.
• Outcomes-based performance measurement is the Holy Grail for professional regulatory bodies. Practically, however, outcomes-based performance measurement is difficult to implement especially in terms of ‘ultimate outcomes.’ Dealing with ‘immediate outcomes’ is more manageable.
• What is agreed on is that there needs to be some kind of causal model which links activities, outputs and outcomes. (The empirical validation of this model is another matter, however).

**Recommendation**

Based on this some of the suggestions in the articles referenced in this report, the idea would be to start with a theory of change (or ‘theory of impact’). From there, a logic model could be elaborated to communicate the model. Finally, a balanced scorecard would be developed to operationalize the model.

```
  Theory of change
     ↓
  Logic model
     ↓
Balanced scorecard
```
Measuring the performance of professional regulatory bodies:
A reading list

Readings with an asterisk could be deemed essential readings.


Professional Standards Authority. (June 2013). *A review conducted for the Royal College of Dental Surgeons of Ontario.*


* Professional Standards Authority. (April 2016). *A review conducted for the College of Registered Nurses of British Columbia.*


