



**PENN PROGRAM ON
REGULATION**

Defining and Assessing Regulatory Excellence

Cary Coglianese
University of Pennsylvania Law School

Discussion Paper for the
Penn Program on Regulation's
International Expert Dialogue on
"Defining and Measuring Regulatory Excellence"

March 19-20, 2015

University of Pennsylvania Law School

Defining and Assessing Regulatory Excellence

Cary Coglianese
University of Pennsylvania Law School

What makes a regulator excellent? This question calls first for articulating a set of *attributes* of excellence, and then for applying those attributes in a process of *assessing* any specific regulator. This brief account details several key issues that arise in both defining and assessing regulatory excellence.

Attributes of Excellence

The first set of issues center on the attributes of excellence. What makes a regulator excellent could be answered in several distinct ways, each of which gives rise to distinct types of attributes.

Characteristics of a regulator (as an organization)

When defining excellence in terms of characteristics, adjectives will be used to describe the qualities or capacities of the regulator as an organization: e.g., “knowledgeable,” “well-funded,” “adequately staffed,” “credible,” “honest,” “legitimate,” and so forth. These characteristics do not describe specific actions or outcomes, although they may well be affected by (or in turn affect) actions and outcomes. Rather, they describe a general “state” of the regulator, a standing set of resources upon which it has to draw or a general posture that it holds in conducting its day-to-day operations and affecting outcomes in the world.

Actions (or best practices) of regulating

Another way to define excellence lies in the type of actions the regulator takes in the course of regulating. Attributes as actions might be articulated in general terms, describing the regulator’s actions in the course of regulating, perhaps again using adjectives such as “vigilant,” “serious,” “reasonable,” “transparent,” and so forth. Or excellence as action might be articulated in terms of specific types of best practices, e.g., “an excellent regulator takes enforcement actions against the biggest risks,” “an excellent regulator uses flexible regulatory instruments,” “an excellent regulator adopts a problem-solving rather than a punitive approach to enforcement.”

Outcomes (or indicia of regulatory performance)

Ultimately the characteristics that define an excellent regulator, as well as the actions that it takes, should lead to desired outcomes. Indeed, what makes certain *characteristics* and *actions* important are ultimately their effects in terms of

helping to achieve desirable *outcomes*. These *outcomes*, then, might be what get used to define regulatory excellence.

Many outcomes, when used as attributes of regulatory excellence, will describe substantive states of the world. For example:

- *effectiveness* (impact in terms of solving the problem or achieve an ultimate outcome of concern);
- *cost-effectiveness* (achieving a specific level of the desired outcome at a low cost);
- *efficiency* (balancing the desired outcome – i.e., problem reduction – with other outcomes or concerns, such as costs, so as to achieve an “optimal” level of problem reduction); or
- *equity* (a fair distribution of the costs and benefits of regulatory action).

All of these examples focus on substantive outcomes. But other outcomes that could be used to define regulatory excellence might be thought of in process-oriented terms. For example, the key attributes of an excellent regulator’s stakeholder engagement process might be defined in terms of legitimacy or trust by the public. If a regulatory process leaves members of the public feeling they were listened to and respected, that is a kind of process outcome that might define regulatory excellence.

Indicia of performance – whether substantive or process-based outcomes – will presumably have some connection with a regulator’s *characteristics* and *actions*. Sometimes this connection will be instrumental, in that a regulator possessing certain characteristics, or a regulator that takes certain kinds of actions, will be more likely to achieve excellent outcomes. For example, a regulator that is highly knowledgeable (a characteristic) will be more likely to achieve effective outcomes. Or as another example, a regulator that adopts flexible rules (action) will be more likely to achieve cost-effective or efficient outcomes. And of course, outcomes may well feedback to shape a regulator’s characteristics or actions too.

Assessing Excellence

Although it is necessary to choose attributes of excellence in order to determine what makes a regulator excellent, the attributes by themselves cannot tell us whether any specific regulator is excellent. The attributes must be deployed in an assessment; in other words, it needs to be determined how well a regulator measures up against those attributes. As a result, a separate set of issues arises when it comes to assessing regulatory excellence. Although some of these issues may seem initially to apply more when thinking of excellence in terms of outcomes, they are actually also relevant to excellence in terms of characteristics and actions.

Measurement

Attributes need to be measured to determine if a regulator possesses more (or less) of a desired characteristic, or whether its actions are usually (or rarely) of the desired kind, or whether most (or few) of its outcomes are of the desirable kind (or how desirable they are). Measurement issues can include:

Metrics. What metrics should be used to measure or “operationalize” the different attributes? What *units* of these metrics should be used? (X alone, or a ratio of X to Y?) Some attributes may lend themselves to seemingly obvious metrics, whereas others may well require more careful consideration. A performance measure for accident reduction may well be an example of the former, whereas attributes such as credibility, honesty, or transparency might be examples of the latter.

Data Sources. What sources of data should be used? Accident reduction might be measured by relying on an established reporting system, for example. Credibility, honesty, or transparency, on the other hand, might need to be operationalized in terms of, say, expert or public surveys that ask others to rate an organization on these qualities.

Triangulation. Should the regulator or evaluator rely on a single metric (or a single data source) for each attribute? Or should the regulator try to “triangulate” and rely on multiple metrics (or data sources)? If multiple measures are used, how should differences or inconsistencies among them be resolved?

Timing. How and when should measurement be conducted? Should measures be taken monthly? Quarterly? Annually? If measuring the effect of specific actions on outcomes, how long should an evaluator wait before seeing if the intended effects are occurring?

Personnel. Who should do the measurement? Should the regulator itself do the measuring? Or should it rely on a third-party?

Aggregation

Under the banner of “aggregation,” a variety of issues can be considered:

Weighting. Presumably no single attribute by itself defines regulatory excellence; an excellent regulator should possess multiple attributes of excellence. Even if one were to choose from only one of the three types of attributes – say, outcomes – presumably there will be more than one attribute that defines excellence: e.g., both cost-effective and equitable. Given the existence of multiple attributes of excellence, one issue concerns the weighting of these several attributes: Does a regulator need to achieve a level of excellence on all of the defining attributes? Most or some of them? Or just one of them? Should different attributes be

weighted differently? Suppose a regulator scores really well in terms of its *characteristics* but it turns out not to generate excellent *outcomes*. Do outcomes count more than qualities? Do some outcomes count more than other outcomes?

Parts. How should differences across different aspects of a regulator's operations (or different parts of its organization) be taken into account? If a regulator's enforcement personnel never take bribes but the officials who set standards do (or vice versa), should the regulatory organization overall be considered to possess a moderate level of the quality of trustworthiness? (Or is an entire fish treated as rotten if just part of it is?)

Combining. A given regulator might well be excellent in terms of some attributes in some aspects of its operations (e.g., setting standards), but not in terms of other attributes on those same aspects. In addition, it might also be excellent in terms of some attributes with respect to another aspect of its operations (e.g., enforcement activities), but not with respect to other attributes. How do the different measures on these different aspects all get combined (if they should at all)? Should the aim be to determine some overall level of excellence? If it is, does each aspect and each attribute count equally? Or are some more important than others?

Attribution

Especially when it comes to outcomes, although also with respect to excellence in terms of qualities and actions, we might consider the extent to which excellence depends on connecting measures of different attributes to the choices made by the regulator. Making these connections is what I mean by *attribution*. There can be both causal attribution as well as attribution of responsibility.

An important kind of attribution is *causal*. Consider the substantive outcome attribute of *effectiveness*. A key question is whether the outcome observed is causally related to the actions of the regulator. Should it matter if improved outcomes in the world are due to the policies and actions of the regulator? Or should the regulator (and any evaluator of that regulator) declare a success and move on, even without finding out why outcomes have improved?

Not every regulatory expert agrees about how to answer these questions. Malcolm Sparrow, for example, has written that “[r]egulatory agencies should not feel obligated to prove causality.” Moreover, because causal attribution is hard to do, it is likely not possible to conduct a causal evaluation for each and every attribute – in which case, the real question becomes one of which attributes should be assessed causally, when and how often.

Although causal attribution may be clearest with respect to outcomes, there are issues of attribution of *responsibility* that can arise with any kind of attribute of excellence. Imagine a regulator that is “doing everything right” but suffers from

low public confidence in its honesty because it is operating within a governmental system that generally has high levels of corruption. Even if that particular regulator itself were to be completely honest, it might be perceived as less trustworthy simply because it is situated within a larger governmental climate that is deeply mistrusted.

Similarly with respect to a regulator's actions, perhaps a regulator fails to follow "best practices" not because it doesn't know better and not because it has chosen not to. Rather, suppose it fails to follow best practices because it operates under a legislative mandate that precludes it from doing so. For example, some people have criticized the U.S. Environmental Protection Agency's air quality standard-setting decisions as being inefficient, but the regulator operates under a legislative mandate that precludes it from taking costs into account in setting air quality standards. Should that legislative prohibition affect an evaluation of EPA's performance in terms of efficiency, as the agency has been told not to consider costs?

In addition to the regulator's mandate, consider other factors that might possibly be out of the regulator's control: e.g., the difficulty of the problems it is tasked to solve, or perhaps a low level of funding available to support its personnel and activities. Is a regulator excellent if it does the very best it possibly can with the limited resources it has been given? If comparing regulators, should it matter if it is easier for one regulator to achieve excellent outcomes because it has been tasked with solving simple problems, as opposed to another regulator who struggles and does the very best that can possibly be done to solve a challenging, even impossible, set of problems?

Or bring back in the issue of the regulator's mandate given by the legislature: suppose the only way to eliminate a problem is to ban a certain economic activity outright, but the legislature has deemed that the relevant private activity should be permitted (albeit regulated). Should the regulator's failure to eliminate the problem count against the regulator in assessing its status as "best in class" or "excellent"?

In cases like these, assessing the regulator calls for connecting the regulator – in terms of responsibility – to the measured attributes when there are other factors clearly at play. Does an excellent regulator in such circumstances have some obligation to try to improve those other factors, especially if they constitute low levels of funding or limited or constrained authority? Perhaps it might be appropriate under some circumstances to attribute some responsibility to a regulator if it fails even to try to secure additional funding or authority.

Grading

As just about any teacher knows, not only must attributes of excellence be identified and measured, but important choices must be made about where and how to establish “cutoffs” for different gradations of excellence. How much of a certain attribute (or combinations of attributes) must a regulator possess to be deemed an “excellent” or “best in class” regulator? Or is it “good enough” just to see improvement over time on the different attributes?

If there is to be a threshold above which a regulator is deemed “excellent” and below which it is not, how should that threshold be determined? Is it to be an *absolute* threshold? For example, should no regulator be deemed excellent in terms of, say, “honesty” if it has more than X level of some measure of corruption? If so, what should X equal?

Alternatively, is the better way to approach this more like with a teacher grading on a curve, namely as a matter of *relative* performance? If so, what is the relevant comparison group? Unlike students taking an exam in a classroom, not every regulator faces the same “exam” or the same exam conditions. Should the reference group be other regulators in the same jurisdiction? Or regulators dealing with “similar” problems in other jurisdictions?

If the last of these is the way to grade a regulator, namely to compare its performance with regulators in other jurisdictions that address similar problems, then issues of attribution of responsibility and causation may become especially important. How should one factor in differences in resource levels, overall governmental capacity, industry characteristics, and demographic, political, geologic, or other environmental conditions -- all which might affect a regulator’s ratings and yet fall outside a regulator’s control?

Decision-Making

Finally, an overarching set of questions encompasses the entire system of performance measurement, and they can be neatly encapsulated in the simple question: Who?

Who picks the attributes? Who selects metrics and decides how they should be operationalized? Who determines how tradeoffs should be resolved in qualities, actions, and outcomes, or in how measures of these attributes should be aggregated across different aspects of a regulator’s operations? In short, whose answers matter most when it comes to all of the questions and issues presented above?

Is it the regulator’s answers that count? Or the answers that are (or would be) given by the legislature or other governmental bodies that oversee the regulator? Or the answers of the overall public (recognizing that there may actually be, in some sense, multiple “publics”)? Do the views of some actors matter more than others? If so, how should these various views be weighted and factored?