Recognizing Public Value: Developing a Public Value Account and a Public Value Scorecard

Mark H. Moore

August, 2012

I. The Challenge of Measuring the Production of Public Value

Creating Public Value left a very large, important question unanswered: how would public managers, the elected representatives of the people who monitored their work, the citizens in whose name public enterprises were carried out, the taxpayers who provided much of the funding for government, and the individuals who had interactions with governments as clients properly determine whether public value was, in fact, being created by a particular government organization?[1] Presumably, the answer to that question lay in the development and use of some kind of accounting scheme that could capture – or, in the accounting sense, recognize – when significant costs to the society were being incurred, and substantial benefits to the society being produced.

A. Utilitarianism

Of course, there had been numerous efforts made in the past to develop such schemes. The most famous was Jeremy Bentham whose book, *Utilitarianism,* was written precisely as a guide to legislators in democratic societies as to how they should judge the public value of specific legislative proposals.[2] The core concept (which eventually became an entire philosophy, not just a practical guide to legislators) was that the social or public value of any proposed public action should begin with individual citizens valuing the consequences of proposed policy on them, and then simply summing the valuations across all individuals.[3] If the sum of individually experienced and evaluated benefits was greater than the sum of individually experienced harms then the policy should be adopted; public value equals the "greatest good for the greatest number."[4]

B. Benefit/Cost Analysis

A century later, social scientists were hard at work to provide the technical means for making precisely the calculation that Bentham had in mind: a calculation of the net social benefit of a given public policy proposed in action.[5] They developed the field of benefit cost analysis in which individuals' "willingness to pay" to secure a benefit or avoid a harm became the conceptual basis for public valuation.[6] Empirical methods were developed that allowed analysts to make these estimates in situations where market transactions would not deliver the desired information directly.[7] A cottage industry grew up of economists and other social scientists who carried out benefit cost analyses of everything from infra-structure projects, through social programs, to the wisdom of investing in a fleet of nuclear submarines.

[1] Mark H. Moore, *Creating Public Value: Strategic Management in Government*
[2] Jeremy Benthan, *Utilitarianism*, opening paragraph
[3] Bernard Williams on Utilitarianism as a philosophy.
[4] This is the core principal of utilitarianism as a philosophy.
[5] Stokey and Zeckhauser on Benefit Cost Analysis
[6] Willingness to Pay Criterion
[7] Methods for measuring willingness to pay

While the conceptual elegance of benefit cost analysis was profoundly appealing to those who believed in the core commitment of liberal democratic theory to the primacy of the individual as the only appropriate arbiter of value in a society, benefit cost analysis ran into a host of practical problems. These included the fact that it was hard to get reliable, objective evidence of what individuals would actually be willing to pay for a particular result in circumstances where they did not really need to pay to secure the desired benefits (or avoid the unwanted harms).[8] But it also included the fact that the empirical estimates about what effects would actually occur in the world as a result of the adoption of a particular policy – the things that individuals were being asked to value – were often highly uncertain.[9] Further, the burdens and benefits of policies were often unevenly distributed over time with the costs coming early and the benefits coming later.[10] Finally, there was the awkward question of what should be done if the distributional effects of a given policy were skewed – with some individuals getting more of the benefits and others paying more of the costs – either in terms of financing or unwanted burdens they were asked to carry.[11]

C. Program Evaluation

To meet some of the practical challenges, a second strand of social science based estimates of social or public value grew up: the practice of program evaluation.[12] This effort emerged more from the field of statistics than the field of economics. Consistent with the rough and ready empirical style of statisticians, program evaluators were prepared to set aside the issue of how individuals valued the particular results (both intended and unintended) of a public policy. They went directly to the core question of what the policy makers must have had in mind when they adopted a particular public policy – the social outcomes that they thought would justify the use of public assets if only they could be produced. Their attention focused on the simple empirical question of whether, and to what degree, the desired outcomes of a public policy had been achieved.[13] To answer that question, all they had to do was to name the specific objectives the policy makers had in mind (or the particular dimensions of public value that legislators and other policy makers sought) and to find or development measurement instruments that would capture empirically the degree to which conditions in the world had moved in the desired direction. Being statisticians, and being particularly concerned about causal attribution, they also spent a lot of time and effort with developing methods that would allow them to make relatively reliable attributions of causation from the observed changes to the changed policy.[14] They used experimental methods when they could, but depended much more often on regression analyses that used natural variation that had occurred to determine whether the policy variables had, in fact, influenced the outcome variables.[15]

[8] Stokey and Zeckhauser
[9] Ibid
[10] Ibid
[11] Ibid
[12] Carol Weiss on Program Evaluation. Harry Hatry on Program Evaluation
[13] Ibid
[14] Causal Attribution in Program Evaluatoin
[15] Experimental and Non-Experimental Methods for inferring causation.

While the empirical contribution of program evaluation to efforts to value public policies was noted and much appreciated as a necessary basis for estimating the effects of public policies that could then be evaluated in financial terms by individuals in creating strong benefit cost analyses, few paid much attention to an implicit conceptual challenge to benefit cost analysis in the routine practices of program evaluators. As noted above, the program evaluation community typically began their analysis by trying to figure out what legislators and other policy-makers seemed to be trying to accomplish through the proposed policy. They thought that the social outcome desired by legislatures or expressed in policy-making processes was the crucially important definition of public value to be used in the evaluation of public policies.

The conceptual challenge to benefit cost analysis in this approach is to reject the idea (perhaps only on practical grounds) that individual citizens in their varied roles might be the appropriate arbiters of value, and to embrace the idea that the proper arbiters of value (or at least those who could be easily found and were articulate) were the legislators or policy makers who set out the purposes that were to be achieved through the use of public assets. Those who wanted to value public policy results didn't have to ask all the individuals what they would be willing to pay; all we had to do was to ask those who represented them what public purposes they were trying to achieve. This simplified the analytic problem of nominating dimensions of public value enormously. But it also changed the conceptual basis of valuation from individual citizens reflecting on their own experience with public policies to representatives of individuals working in the context of democratic government to speak for a collective about what was valuable. From the point of view of democratic theory, one could argue that this was not simply a matter of convenience, but a crucial correction in the theory of benefit cost analysis which imagined that the appropriate arbiter of public value were individuals, not the public as it was constituted through democratic politics.

This also likely changed the substantive form of what was valued. Instead of trying to satisfy individual utility functions (assumed to be focused primarily on an individual's own material self-interest), legislators and policy makers in the executive branch sought to define the purposes or mission of a government enterprise. This takes the form of a social utility function that is formally and substantively different from the satisfaction of individual material self-interest.[16] The arbiter of value is a collective formed through democratic processes, not an individual. And that collective arbiter of value might well be interested in aggregate social results including distributional effects as well as the simple sum of individual valuations.

Note that this way of valuing the results of public policies is still a utilitarian approach. We have not given any special standing to the idea that individuals might have rights as well as interests, and that a valuable feature of any public policy might be the degree to which it seemed to undermine or actually violate individual rights. Nor have we explicitly introduced philosophical concerns about justice and fairness in the ways that public policies work, and the results they achieve as an important consideration in valuing or evaluating public policies. So, the analysis remains comfortably in the realm of utilitarianism. But it is not a utilitarianism in which individuals are the arbiters of value. Nor is it a

[16] Stokey and Zeckhauser

utilitarianism in which all important values can be easily measured in financial terms. This is important to remember as we go forward, because one of the important challenges we face in figuring out how to value public policies is to the degree to which we should depart to some degree from utilitarian thinking and incorporate concerns for rights, justice and fairness as well.

So, while it is common to view program evaluation as a kind of poor cousin to benefit cost analysis (which remains the gold standard for the proper evaluation of public policies), one can also see program evaluation as a conceptual challenge to the central focus that methods of benefit cost analysis give to individual valuations, and also to the monetization of all relevant effects of a policy. Program evaluation suggests that there might be a way of evaluating government programs which depends on values assigned by a collective judgment created through democratic processes not by individuals; further, that the desired effects could include results that could not easily be monetized, but still could be objectively noted and recorded.

D. The Business Sector's "Bottom Line"

Interestingly, none of these methods looked much like the methods that are used in business, or more generally in management, to guide organizations towards improved performance.[17] As far as I know, businesses do not carry out benefit cost analyses. Nor do they carry out program evaluations. What they do do is to assess the profitability of organizational activities.[18] They do so by comparing the revenues earned by the sale of products and services to customers who chose to spend their money on those products and services with the costs the organization had to lay out to produce the goods and services it offered for sale. Sometimes the unit of such an analysis is the organization as a whole, sometimes it is a strategic business unit within the organization, sometimes it is a particular product line or line of service, other times it is a particular production process that is used to produce specific products and services. The only information system the organization needs to support this financial assessment is a financial system that keeps track of costs paid out, and revenues coming in from paying customers. Since all companies have such systems in place (at least in part to avoid theft, and in part to meet reporting requirements set by government), it is not hard for them to use bottom line performance evaluations to assess performance at many different levels of aggregation in the organization, and to use such assessments to increase pressure to improve performance, and to make better resource allocation decisions. [19]

This is such a commonplace occurrence that it is easy to overlook how important this financial system is in helping private sector managers manage their organizations for value creation, and what

[17] Strategy and Control Processes in Business. Bower, The Resource Allocation Process. Kaplan and Norton, The Balanced Scorecard.
[18] Business literature on strategic business units and the use of financial bottom lines to allocate resources.
[19] It is enlightening to think about the history of the development of financial accounting systems. They were initially developed by monks who wished to keep books to show that none of God's bounty has been diverted or carelessly used or stolen. They could account for its use. It took a few centuries for businesses to learn how to use the accounting systems originally designed to prevent theft to determine whether and how their organizations were producing value for customers and shareholders, and to develop systems for doing this consistently across all organizations. Perhaps we are in the early stages of this process in accounting for the creation of public value.

exactly is missing when we try to bring the same concept of bottom line management into the public sector. In principle, public sector accounting systems can be as good as private sector systems in capturing the most obvious material costs of producing desired social results.[20] Government organizations rely on inputs of labor and materials to produce results. They often buy these in open markets, and pay market prices for the inputs. The costs can be assigned to particular organizational units and activities. In short, there is very little difference between private and public sector cost accounting.

What is missing from the public sector accounting system is the functional equivalent of the revenue measure that the private sector relies on. It is not that the public sector doesn't have financial revenues to account for. Obviously money flows to the government in the form of taxes, and those funds are used to sustain government activities. What is different is the *philosophical meaning* of the revenues used by private and public organizations. In the private sector, when an individual plunks down their hard earned money to buy a particular good or service, we can reasonably conclude that that individual valued the good or service they purchased. We even have a pretty accurate estimate of precisely how *much* they valued that good or service! We can also make a direct comparison of the value that individuals attach to particular goods and services and the cost of producing them simply by comparing the revenues earned to the costs of production. If that comparison shows that the producing company has earned more revenues than it has paid in costs – that it is "profitable" – then those who manage the company can assume not only that they can pay their investors and survive another day, but also that some kind of social value has been created. If a liberal society thinks that individual valuations are important arbiters of social value, and if we have direct evidence that individuals are willing to pay for a particular product or service at a price that covers the cost of production, then we can have some faith that something of value is created in profitable companies.[21]

This reasoning is much less tight when government enterprises are financed not by individuals purchasing goods and services for themselves at a point of sale, but instead through taxes that are imposed on them as their fair share of financing an enterprise that a collective has judged to be important. Revenues are less directly attached to the satisfaction of individual citizens, taxpayers and clients, and less directly attached to the performance of government organizations.

This is all well known, of course. But the full consequences of this fact for efforts to value the performance of government organizations has not yet been fully appreciated – witness the continued confused discussion about the customers of government.[22] To get a sense for the magnitude of the problem created by the fact that government enterprises do not have a revenue number tied to the voluntary choices of individuals to buy the product, all one has to do is to consider what private sector investors and managers would do if they, like the government, could have all the cost information they needed, but no record of revenues earned through sales. Because revenues earned from the sale of

[20] I am temporarily ignoring here problems I raise in my philosophical paper about the inputs of authority and public spirit.

[21] Many challenges to this simple story. But it is a pretty strong simple story, and one that carries great weight in these commercial times.

[22] Customers of government debate.

products and services form not only the practical goal of a private enterprise, but also their social *raison d'etre* – the demonstration that they have made good use of the resources they have claimed – if they lose the measure of revenue, they can no longer know whether they are producing value or not. There is no market test giving them the feedback they need as to the social value of what they producing.

Viewed from this perspective, we can see that benefit cost analysis and program evaluation were trying to do is to create the functional equivalent of the revenue measure in the private sector: an objective basis for determining whether and how much value in the form of socially desired results was actually being produced by government. The most obvious difference between a measure of revenue earned through the sale of products and services to individuals who decide individually that they want and are willing to pay for a particular good and service on one hand, and benefit cost analysis and program evaluation on the other is how much more difficult it is to create the public measures of value, and how much less reliable they seem when they are created.

One important implication is that many fewer things can be evaluated in the public sector than in the private sector, because the set-up for the evaluation is so much more cumbersome. That means that, for any given administrative cost in developing and using information and evaluation systems, public managers can probe less deeply and less broadly into the value of their activities than private sector managers.

But one should also notice two other big differences between the private sector's revenue measure on one hand, and the public sector's benefit cost analysis and program evaluation on the other. First, as noted above, the private sector's revenue number treats the individual and his experience as a customer of the organization as the important measure of social value. While benefit cost analysis seeks to hold onto the idea of individual valuation (ideally in financial terms), program evaluation often departs from this idea and treats the important arbiter of value as a public that is formed through government, and the results that are valued as potentially different from what individuals desire for themselves. Similarly, the private sector's revenue number captures the value of a good or service to individuals in financial terms. While benefit cost analysis often tried to go beyond the concrete effects of a given public policy to monetize the value of those effects in the minds of individuals, program evaluation once again departs from this model by contenting itself with the simple measurement of effects, without trying to monetize them.

E. Management Accounting in Government

Of course, government agencies could not escape the demand for organizational accountability. They needed some kind of system to report to their "investors" and "shareholders" about their performance, and to manage their operations for efficiency and effectiveness. This was true even if they found it difficult to account for the value they were creating!

Their response was to create accounting and measurement systems that monitored their control over assets, compliance with policies and procedures, and organizational outputs and transactions that occurred at the boundary of the organization. Internal audits sought to ensure that government money was not be stolen or diverted to unauthorized purposes.[23] Compliance audits sought to ensure that government assets were being used according to established policies and procedures.[24] Outputs of organizations were counted to measure productivity and workload. Etc.[25]

This was all fine as far as it went. It was important to ensure citizens and taxpayers that tax dollars were not being diverted through official fraud, waste and abuse. Indeed, insofar as many government programs involved little more than the distribution of financial resources from the Treasury to individual client beneficiaries of government programs, ensuring that government dollars went to individuals who were the intended beneficiaries and not to others went a pretty fair distance in ensuring that the program was creating the value legislators and policy-makers intended.

Moreover, insisting on compliance with existing policies and procedures was valuable was potentially valuable for two slightly different reasons. On one hand, since treating like cases alike was considered an intrinsically valuable characteristic of government activity, and since compliance with policies guaranteed such consistency, one could say that compliance audits produced a much desired fairness in government operations. On the other, if existing policies and procedures embodied the best possible methods for using government assets to achieve desired results – if they encoded what was known to be best practices in producing valuable results – then compliance with those policies would ensure efficiency and effectiveness as well as fairness. [26]

The problem was, of course, that these accounting systems did not reach far enough down what could be described as the value chain that stretched from government controlled assets, through organizational activities, to outputs of government agencies, that could then be valued by clients who encountered the government, or by society as a whole in the achievement of socially desired outcomes.[27] These systems stopped at the organizational boundaries -- well before it was possible to see the ultimate value that was being produced. That left citizens, taxpayers, and their representatives uncertain about what had actually been accomplished. They knew how assets were being deployed, and could see the transactions with clients at the business end of the government agencies. But they did not know how to value the effects of those transactions. The revenue measure that the private sector relied on so heavily for evaluating its performance was missing.

Efforts to fill this gap with benefit cost analyses and program evaluation could not be entirely satisfactory for at least the following reasons. First, as noted above, these analyses were often quite expensive, and difficult to organize. That meant that only some would be done, and that there would be many government activities which could not be evaluated with these methods. Second, the results of these analyses tended to come in infrequently and late. This made them potentially important in making big policy decisions, but hard to use for managing operations – including adapting operations in *media res* when they did not seem to be working. Third, these methods were often focused on the evaluation

[23] Financial Audits in Government
[24] Compliance Audits in Government
[25] Workload and Output Measurement in Government
[26] Danger of best practices that aren't
[27] For picture of the value chain, see Figure ____

of policies and programs – not organizational units. Of course, this was not a problem if policies and programs were uniquely identified with particular organizational units. But if a policy cut across several different organizations, or if a program was only part of what one organizational unit did, then these systems would, once again, be useful in policy and program decisions, but not as useful as business' bottom line in the management of organizational resource allocation, motivation, or learning.

Finally, it remained true that all of these efforts to measure the social or public value of government operations lay within the utilitarian framework. This is not surprising given the practical, business oriented world in which we live. But government could not quite insulate itself from continuing public concerns about whether government was improperly intruding on individual rights, whether it was treating those who it regulated and to whom it provided benefits fairly, and whether it was helping to produce a civil and just society, as well as one that was prosperous and good. These concerns that will not go away are linked less to utilitarianism and more to what political philosophy would describe as deontological concerns: questions about what we owe to one another and the quality of our relationships as to our individual material satisfaction.[28] If we wanted a government that acted fairly and justly, and helped to encourage fairness and justice in the society as a whole, some method for incorporating these concepts into the measurement of public sector operations would have to be found.

F. Kaplan's Balanced Scorecard As Challenge and Opportunity

In 1996, the world of performance measurement -- both private and public – was rocked by the publication of *The Balanced Scorecard*.[29] This book was intended principally for private sector managers. It argued persuasively that private managers had long over-estimated the value of their financial systems to guide their organizations. They did not say that the goal of maximizing long term profitability should not be the goal of private sector organizations.[30] Nor did they raise any doubts about the adequacy of financial measures to fully capture the value of what had been produced by private organizations, or what would be produced in the future.[31] In short, they were not recommending the use of double, triple or quadruple bottom lines for private sector organizations. [32]

What they were saying is that the financial measures used by private sector organizations were always about the past.[33] They captured what the organization had done in previous years. They told the organization very little about what they needed to do now to sustain profitability in the future. And from their point of view, that was what all the important business decisions were about – how could profitability be sustained in the future?

To answer that question, business executive had to divert their attention from backward looking financial performance measures, and look into the future and develop a plan for future profitability –

[28] Frankena, Ethics.
[29] Robert S. Kaplan and David P. Norton, *The Balanced Scorecard: Translating Strategy into Action* (Cambridge, Mass. : Harvard Business School Press, 1996)
[30] Ibid
[31] Ibid
[32] Multiple Bottom Lines for Business Accounting. Sanjeev Khagram, Jane Nelson, Case on Ethos Institute
[33] Kaplan and Norton

what they described as the business strategy of the organization.[34] To do that, business executives had to have good information about their market position – particularly their standing with customers. [35] They also had to have a sense of whether current operations could be sustained, and how investments could be made to change what the organization was producing or how it was producing the particular goods and services it offered for sale.[36] This led them to propose a performance measurement system that embraced four perspectives.[37] The first was the financial perspective which still occupied pride of place as an accurate description of the value to be produced by the organization. But then they added three more perspectives: the customer perspective (what is the state of our relationship with customers whom we will rely on in the future to buy our products and services?); the operations perspective (how efficient are our current operations and can they be sustained?); and the learning perspective (how can we adapt both our products and services, and the methods we use to produce them to be more profitable in the future?).

This did not come as big news to business executives who had long deployed many measures other than financial measures to monitor current operations, and to guide investments for future oriented strategy execution. But it did serve to emphasize the importance of all the following kinds of measures above and beyond financial measurement in private sector organizations: 1) non financial measures; 2) measures of the external environment rather than what was happening in the organization and at its boundaries; and 3) process measures that were focused both on internal operations, and the investments being made to transform current product lines or operations. Insofar as the ideas began to undermine the overwhelming dominance of the financial bottom line as the only way to capture the value or an organization's operations, or manage it for future performance, *The Balanced Scorecard* deeply challenged and re-ordered private sector measurement practices.

To managers in the non-profit and government world, *The Balanced Scorecard* seemed as much an opportunity as a challenge. Here was a respected business expert explaining why financial measures could not be the only measures used to value and guide organizational activities. High performing organizations needed to have non-financial measures as well as financial measures. They needed measures that were aligned with an explicit theory of value creation not just a demonstration of past accomplishment. They needed measures focusing on internal organizational processes and investments as well as ultimate outcomes. That was exactly what they had long been saying about why financial measures were not necessarily appropriate for organizations whose revenues came from third party payers such as government and charitable donors rather than from paying customers. Finally, they had someone who could champion their cause! And, again a cottage industry grew up that applied the ideas of *The Balanced Scorecard* to organizations in the non-profit and government world.

What nonprofit and public managers who tried to apply these measures found, however, was that *The Balanced Scorecard* left unanswered two critical questions about performance measurement in government organizations. The first was the old bugaboo: who exactly should be viewed as the arbiter

[34] Ibid
[35] Ibid
[36] Ibid
[37] Ibid

of the value produced by such organizations, and what exactly was it that they either did, or should value? Kaplan and Norton's book directed organizations to identify their customers. But confusion reigned about who the customers of nonprofit and government organizations were? Were they the folks who paid for the production of goods and services? (In which case, one would say that the customers of nonprofit organizations were the donors, and of government organizations, the taxpayers.) Or, were they the folks who were supposed to benefit from the organization's activities, and met the organization in individual transactions. (In which case, one would say that the customers of the organization were the client/beneficiaries.)

One might say that this shouldn't matter, but it did! There was no guarantee that taxpayers wanted the same things as government beneficiaries. In fact, it seemed likely that their interests were opposed. Similarly, there was no guarantee that donors wanted the same thing as their beneficiaries. And if there was a difference, which of these different customers would be properly viewed as the morally compelling or practically important arbiter of value?

The second was that, once again, the accounting scheme being developed for the public sector was principally an expression of utilitarianism – a way of looking at costs incurred, benefits produced without much attention to questions about the kinds of relationships that donors ought to have to beneficiaries, or government to citizens as a matter of justice. Nor was there any attempt to explain or to show the relationship of a non-profit or governmental organization's activities and the achievement of any particular kind of social solidarity or justice. Since much of the justification for both charitable activity and government activity seemed animated by individually or collectively held views of what was good for others, and an ideal of a good and just society, it seemed odd not to be able to capture that aspiration and that value in the accounting scheme that sought to recognize the value produced.

II. Recognizing Public Value: A Strategic Approach to Performance Measurement in Government Enterprises

In *Recognizing Public Value*, I have tried to find a different path for the future development of measures that could: 1) more or less reliably capture the public value being produced by a government enterprise; 2) meet public demands for accountability in government; and 3) help managers use performance measurement systems to guide their organizations towards improved performance.[38] This different path consists of three different elements.

A. Taking a Developmental, Strategic Approach

The first is to recognize that, while the development of suitable performance measures for government might eventually simplify oversight and management in the public sector, in the short run it is likely to create some increased complexity. To many, this will come as unwelcome news. After all, part of the appeal of "bottom line management" for government is precisely that it might make things relatively simple and objective. And it can't really be that hard. Claims that it is difficult to do can be easily seen as another form of bureaucratic obfuscation designed to prevent any kind of real

[38] Mark H. Moore, Recognizing Public Value, Forthcoming

accountability. All we have to do is be tough and determined in developing and using the measures. It is time to get on with it. Let's spend less time arguing about what should be produced, we can get on with producing it.

I am all in favor of being tough and determined in the pursuit of political and bureaucratic accountability through the development and use of quantitative, objective, performance measures. But I am painfully aware of how much can be lost both in organizational performance and in the cause of performance management if a large commitment is made to the wrong set of measures. Indeed, I think one can walk into any public organization and find in their organizational structures, and in their information systems, the rusty hulks of previous efforts to measure performance that have been abandoned by those who once created them with great enthusiasm. Of course, one could say that the reason these systems have been abandoned is that the bureaucrats finally wore down the politicians' and managers' efforts to impose a reasonable kind of accountability. But I think a bigger part of the problem is that the measurement systems that were created were not very good systems. In a surge of enthusiasm for performance measurement, organizations reached for a set of measures that could be quickly used without really worrying too much: 1) whether the measures could reliably capture the value that the organization was producing; or 2) whether the measures were aligned with the values that those in a position to call the organization to account wanted to see achieved by and expressed through the organization's operations; or 3) whether the measures could attract the loyalty and commitment of those working within the organization who were being asked to labor on their behalf.

An alternative approach would recognize that the development and use of a performance measurement system was like the construction of a cathedral: even though it could inspire and to some degree command when it got to a certain point, it was never quite finished. There was a lot of work in the initial construction. But however much work it took to build the foundation, anchor the buttresses, and build the spire, more work would be necessary to create the embellishments; even more to make the significant renovations as conditions change; and still more to re-build when the cathedras was razed by barbarians! Just as a firm's *Balanced Scorecard* might change in response to environmental changes, and to new strategies to deal with that change, so the performance measures used by government policy makers and organizational leaders might change with the times. The measures could remain objective, but they are not eternal.

I call this alternative a strategic approach for two different reasons. First, I want to emphasize that there is an important developmental aspect to the creation and use of performance measures in government. Part of that development is doing the basic investment work to create a plausibly good system now, ignoring the temptation to limit oneself to what is currently available in the organization. Another part is anticipating that after one has built a pretty good system, it will probably have to change in important ways to accommodate changes in the world. Measurement is an instrument of a strategic calculation about how to best use the assets held within an organization to create public value in the world in which the public organization is operating. It is not an objective thing that is fixed and eternal.

Second, I want to emphasize that a valuable performance measurement and management system would be one that played an important integrating role in fitting together the pieces of a strong strategic design in the public sector. In *Creating Public Value*, I argued that a good, value creating strategy for a government organization or enterprise had to meet three tests: 1) it had to make a plausible claim that the envisioned purposes were publicly valuable; 2) it had to be able to command legitimacy and support from those who authorized and financed the activity; and 3) it had to be operationally achievable. The challenge was to integrate these pieces in the short run, and to envision how to move to an organization or enterprise to an improved position in the future. This meant that government managers had to have not only measures of ultimate outcomes and value, but also, like the private sector managers who were using the *Balanced Scorecard*, to have some way of monitoring their current position and capacities, and to envision and execute strategies that would sustain or improve their position and performance in the future.

To no small degree, private sector managers take up their work leading organizations with much of the work of building a performance measurement system already done. They have the financial measurement system, and the human systems that are comfortable with using these systems for both external and internal accountability. Public managers, however, have both *more* and *different kinds* of work to do to create a measurement system that can do the work that the financial measurement system does in the private sector.

B. Four Different Kinds of Managerial Work

In *Recognizing Public Value,* I claim that there are four different kinds of managerial work to be done in creating a valuable performance measurement system. I describe them as technical, managerial, philosophical, and political. The first two are familiar and widely acknowledged but still difficult. The last two others are less widely recognized as important, but, in my view, absolutely crucial to the development of a value creating, sustainable measurement system. Indeed, I think part of the reason we have not done very well in developing performance measures is that we have concentrated on the first two, and neglected the second two.

By technical work, I refer to the challenge of developing specific measurement instruments that can accurately capture the degree to which valued (and disvalued) effects are occurring in the world while the enterprise is operating. An accounting system has to start with some ideas about the kinds of effects that could occur in the world, and which, if they occurred would register in some agent's idea of whether that effect was valuable or not. Once one has a concept that defines a particular public value, or a particular dimension of public value, that an enterprise is supposed to produce, someone has to develop some technical means for determining whether and to what degree that effect is occurring in the empirical world. One could say, for example, that an urban police department's value creating mission was to reduce crime and promote security. Assuming this claim commanded widespread political agreement and had a strong philosophical basis, the technical challenge would be to find a way to measure the degree to which these conditions were being advanced in the world. We might develop a measure of reported crime and see whether reported crimes were going up or down. But we might want to supplement that with victimization surveys that tried to get at the "dark figure" of crimes that went unreported to the police, or that sought to render objective the subjective experience of fear of crime. These are examples of technical questions of measurement.

By managerial work, I mean the actions that managers have to take to build the performance measures, and to use them in their organizations to drive performance. To many, using performance measures to animate improved performance is a relatively straightforward task. All one has to do is to attach specific performance measures to particular managers or workers, and use them to reward good performance and punish bad, and the performance of the organization will improve – at least in terms of the measured dimensions of performance. But as many managers can attest, creating and using a performance system that an organization can tolerate as a fair and appropriate system, or that can engage its managers and workers in a whole-hearted effort to improve performance (rather than figure out ways to cheat or game the system) is not as straight-forward and simple as it seems. Things get even more complicated if, instead of simply driving individuals to perform better against the performance measures, the managers decide that they want to help the organization *learn* about what works rather than simply work harder. The information and social relations required to create an organization that has a passionate commitment to doing its job well, and continuously learning how to do it better, are harder to construct than a simple strict liability system that imposes sanctions on employees on the basis of statistically viewed performance.

There is much to occupy managers trying to develop and use performance measurement systems in the two areas set out above. And many efforts fail because of failures in one or the other of these domains. But the two other domains of work – philosophical and political – must also be managed.

By philosophical work, I mean the work that has to precede the technical work described above. The idea of social or public value is a normative, philosophical concept. So is the idea of performance. We can and do try to make the ideas of value and performance more objective and technical by constructing specific empirical measures that we hope can stand reasonably well for the normative ideas. But it is important to recognize the all performance measurement and management systems have at their core a normative theory of what would be valuable and what would be costly if an organization produced those valued or disvalued effects.

In the discussion above about the strengths and limitations of other approaches to measuring performance, we made reference to two quite different philosophical traditions in defining value.[39] On one hand, utilitarianism locates value in material conditions that occur as the consequences of action. We distinguished between individual utilitarianism where the evaluator of the conditions is an individual, and the object of the evaluation is how their material well being is affected; and a more social or collective utilitarianism where the evaluator of the conditions is a collective, and the object of their consideration is the material conditions of the society as a whole. On the other, deontological traditions find value in the degree to which individuals live up to their duties and responsibilities to one another, and the degree to which institutions behave as they are supposed to behave. This is a more legal or juridical idea. The normative language used involves concepts of fairness, or justice, or right relationships, or duties and responsibilities as well as rights and privileges.

---

[39] FRankena, Ethics

As I will argue below, all public sector organizations have to be concerned with evaluating their activities in both philosophical schemes. Those who gain or lose from government operations, and those who are treated fairly or unfairly, have interests in how public organizations conduct themselves, and will want to be able to talk about the value created by government organizations in terms that come from both philosophical traditions. Because those in a position to value the performance of government organizations value these different dimensions of performance, both would have to be included in a scheme that sought to capture the value produced by a public organization. To use the example of policing again, citizens will be interested not only in whether the police are successful in reducing crime and fear at relatively low cost, but also whether the police treat individuals fairly, and whether they help to create justice in the society as a whole.

By far the most challenging kind of work in developing a performance measurement and management system in government is the political work that is involved. This work is crucial, of course, because in a democratic society and a democratic government it is not enough to have a system that is philosophically and technically sound, and managerially useful; it also has to be endorsed and embraced by those in positions to authorize, legitimate, and pay for the enterprise. This is *morally* important because the only appropriate arbiters of public value in democratic political systems are citizens, taxpayers and their elected representatives. It is *practically* important since it is the political demands for accountability that create the drive to create and energy to develop and use performance measurement system. Without political agreement and commitment to a particular performance measurement system its effects will be as dim as a lamp that is not plugged into an electric socket. In contrast, with political support and interest, the performance measurement system not only clarifies the mandate for government operations, it focuses and sustains the energy through the difficult implementation phase of public initiatives.

At the outset, it might seem surprising that one would have to worry about the political demand for accountability. It often seems that there is plenty of that to go around. And that is certainly true. What I have elsewhere described as the political authorizing environment for public agencies is full of those who can be usefully described as “accountability agents:” social actors who feel that they have at least the right and sometimes the responsibility to call public organizations to account for some aspect of their performance. It is the set of these accountability agents, and the particular dimensions of organizational performance that concern them that create the context of social accountability for organizations, and the arena within which public organizations have to fight for legitimacy and support.

Some of these accountability agents occupy positions in government that give them the formal right and responsibility to oversee public organizations, to call them to account for their performance, and make suggestions for improvement. Others – perhaps most – are self-appointed to the job, and no matter how partial or idiosyncratic their interests, they persist in believing that their particular demands were fully aligned and consistent with the public value the organization was supposed to produce. Figure 1 presents a picture of agents who can call the NYPD to account, and identifies the particular dimensions of performance that the particular agents monitor.

[Insert Figure 1 Here]

The implication of this observation is that the problem that has to be solved in the political world is not the *amount* of accountability that is demanded; it is the persistence, the coherence, the focus, and the adaptability of the cumulative demands for accountability.

Take first, the issue of persistence. It might seem surprising that one would have to worry about whether accountability agents would show up to demand accountability from public enterprises, and take the enterprise seriously – particularly when there are officials appointed specifically to this task. But there is lots of evidence to suggest that oversight of government operations by elected officials in both legislative and executive branches is a bit spotty. Their interests rarely run across the whole of government operations, but are instead focused on particular bits in which they have a particular interest, or which have become notorious for some reason. Moreover, even the partial interest they take is apt to be temporary – lasting only as long as their interest or the notoriety that first drew them to the issue. What is most important to such officials is that they appear to be tough in demanding accountability. And they do that by pounding their fist about very specific instances of government malfeasance.

Given this, the first order of business in setting up a performance measurement system would be to see if an agreement could be negotiated between overseers and managers about the specific terms in which the managers will be called to account. If such a deal could be made, the managers could go forward to try to produce the desired results. Such a deal would seem to be in the interest of political overseers since it is consistent with the discharge of their duties and helpful in guiding government organizations towards a conception of public value that is legitimated by their electoral standing, and the quality of their deliberations. Yet, in reality, few political overseers are willing to commit themselves to such a deal.[40] They are reluctant to give up the discretion they would otherwise have to decide what issues they would like to take up with an organization. Since their public visibility often depends on finding and exploiting issues about government performance, it asks a lot of them to be willing to limit their interests in advance to only a certain number of discussable issues. They would like to be able to range freely over the variety of issues that could be made about an organization's performance.

The consequence of that for managers, however, is that they never know what to expect, and find it hard to concentrate on improving performance across the board. Instead of facing consistent strong pressure to improve accurately measured performance on important dimensions of value, they face the prospect of lightening striking in the form of an unexpected issue – and perhaps a trivial one to boot. Their natural response is to turn away from transparency of their operations on a limited number of issues to secrecy. With that, public accountability is lost with adverse consequences for both the performance of the organization, and the maintenance of a proper fiduciary relationship.

Take, next, the issue of coherence. Coherence would seem to be an important feature of a structure of accountability. If one is called to account for producing things that seem inconsistent with one another, it is confusing and demoralizing. It is hard to know how to succeed.

[40] See case and discussion in Chapter 3 of Recognizing Public Value

Yet, it is important to note that the concept of coherence can refer to a psychological state (how easy it is for managers and those overseeing their work to see an alignment among different assignments and responsibilities assigned to them), or to a technical reality (the degree to which the pursuit of one goal or valued attribute of performance helps to advance the other goals as an empirical matter). Initially, it is easy to see that managers confronting their anarchic demands for accountability will experience their situation as incoherent, and therefore frustrating and impossible. They long for their situation to be resolved by the creation of one measure that can capture all, or a concept formulated at a high enough level of abstraction that it can appear to dissolve all the more particular, potentially competing concerns in its warm embrace. And this is what often happens.

But this response is primarily to deal with the psychological discomfort of trying to produce valuable results on many different, and potentially competing objectives. What should be important to managers and their overseers is to what extent the different objectives are really consistent or inconsistent with one another – or more precisely, how much of one valued purpose we might have to give up if we thought another purpose was also valuable, and the actions being contemplated affected both dimensions of value.

This asks managers and their overseers to understand what economists would call the production possibility frontier that they face in deploying the assets of an organization.[41] We may want a police department to control crime, reduce fear, treat individuals fairly and justly, and do so at a low financial cost to the city. Common ideological views make it easy for us to assume that there is an important "trade-off" among these competing values: that we will have to give up some crime reduction potential to advance the goal of keeping costs low, and protecting individual rights even as we are trying to call offenders to account. And that may well be true. But what economists also tell us that we have to face these trade-offs among competing values *only when organizations are operating at the production possibility frontier* – that is when they have exhausted all opportunities for improving on all dimensions simultaneously. If organizations are operating *inside* the production possibility frontier, there is a chance as an empirical matter that we can improve in all dimensions simultaneously. Contrary to all common sense, there is such a thing as a free lunch if we are smart enough to see it and exploit it.

In the private sector, we often assume that private firms are operating at the production possibility frontier – that they have exhausted all opportunities to improve their performance in controlling cost and improving the quantity and quality of output. We assume that because we assume that competitive pressures in an industry have forced all firms that survive to operate on this frontier. Yet, nearly always when private sector firms are challenged to improve their performance – for example, when US auto firms were challenged to meet competition from Japanese firms on both cost and quality – they find that there are many ways to improve performance on both dimensions.[42] They were not, in fact, on the production possibility frontier, and did not know this until they were forced to consider this possibility.

[41] See discussion of production possibility frontier in Chapter 7 of Recognizing Public Value
[42] Quality is Free Article in HBR

If this is true for private sector firms where there are pressures forcing them to the production possibility frontier, how much more likely is it that a public sector organization without equivalent pressures to perform is on the production possibility frontier? If public firms are not on the production possibility frontier, then there ought to be room to improve on all dimensions. Instead of agonizing over the painful choice of which dimension of performance we value more than others, then, we could just assume (until we are proven wrong empirically not ideologically) that we can find a way to improve performance on all dimensions simultaneously. And we can use a performance measurement system that monitors performance on all dimensions of value to show us where that production possibility frontier lies, and how we might move towards it.

From this perspective, then, the concern about coherence in the accountability system might be overblown. It is important to remember that the core challenge facing private managers – to maximize profitability – is itself a representation of dimensions of value that are inconsistent with one another. On one hand, in order to earn revenues, we have to spend money to produce products and services. On the other, to earn profits, we have to reduce costs. So, should we try to maximize revenues or minimize costs? The answer, of course, is that managers try to do both, and measure the value of their efforts that relates the two dimensions of performance in a simple equation: Profits = Revenues – Costs. Perhaps the solution to the technical (if not the psychological and ideological problems of incoherence) is for us to name all the dimensions of value that are considered important in evaluating the performance of a governmental enterprise, remind ourselves of which of these we value positively and negatively, and begin looking closely at our real experience in producing along these different dimensions.

Take next the issue of the focus of demands for accountability. By focus of accountability, I mean two things. First, where along the value chain that causally connects a flow of assets controlled by an organization, through a set or policies, programs and activities carried out by the organization, to outputs of the organization, to transactions with clients of the organization, and ultimately to social outcomes produced by the organization. In principal, external, political accountability systems could be focused anywhere along this chain. (See Figure 2 to illustrate both the value chain, and the instruments used to monitor performance at each stage.) Second, whether the focus of accountability is on individuals and individual incidents, or on more aggregate patterns representations of performance.

[Insert Figure 2 About Here]

In the past, many of the formal systems of government accountability focused heavily on aggregate measures of inputs, policies, and outputs. As noted above, recent trends have encouraged government agencies to focus measurement and accountability further down this causal chain to the loci where public value is being created – in the quality of the transactions with individual clients, and the degree to which desired social outcomes were being achieved. But government organizations have long been called to account not for aggregate patterns, but also for individual events. On one hand, the media have always feasted on individual incidents to create questions about the performance of the larger system. On the other, both courts and proliferating citizen complaint bureaus have given individual citizens and clients the chance to call agencies to account for individual transgressions. And, as noted above, elected politicians have always felt comfortable to raise issues across the broad surface defined by stages in the value chain, and both aggregate statistics and notoriouos incidents.

Again, it would be nice if this system of accountability could be focused a bit more consistently. For example, many would emphasize focusing on social outcomes or aggregate measures of individual client satisfaction as the best measures of performance . That is what is often meant by bottom line management. And many think that would be much superior to a system that focused on process, or that encouraged political overseers to micro manage and give undue attention to incidents that were not really reflective of the usual performance of an organization. But it seems unlikely that this is where the answer lies. It is certainly true that driving down the value chain towards client satisfaction and the achievement of social outcomes represented an advance in our ability to recognize public value when it was being produced. But it is important to remember that there are many features of what is called process that are valued intrinsically by citizens and their representatives as well as instrumentally in terms of their ability to produce desired results. Indeed, these turn out to be particularly important when the task is to recognize the performance of government organizations that are in the business of delivering obligations to citizens rather than benefits, or to ensure that benefits flow to individuals who are entitled to them and not to others. In these cases, the individuals witnessing the encounters as well as those experiencing them will want to know that they have been treated decently and fairly. So, in all likelihood, measurement system designed to meet external political demands for accountability will have to continue to respond to demands across this surface: managers will be called to account for process as well as outcome, and for individual incidents as well as aggregate patters observed at each step along the way to public value creation.

Consider finally, the question of adaptability. To some extent, a virtue of measurement systems designed to meet external political demands for accountability is that they are not adaptable. They should be relatively rigid to command attention, to allow the accumulation of experience over time, and to guard against political manipulation. Yet, taking a strategic approach to the development of these systems means that we have to recognize that they will change over time. This can happen at the technical level as we learn more about the technical properties of the measures we are relying on to connect philosophically and politically valued dimensions of performance to events in the real world. It can happen as managers learn more about how to use the systems to generate internal accountability for performance and continuous learning. And it can happen in the political world as political aspirations change, and the specific content of the values to be advanced through and reflected in agency operations changes in response.

To many, the problem is not the accountability system does not adapt to changing political aspirations and desires; it is that it changes too often – that the political winds that sweep over the agencies are too fickle making it difficult to maintain a steady course. And, viewed from the perspective of several years, that may well seem to be true. The important dimension of value to be pursued in the first year of an administration, may be different than in the fourth year, and different still when a new administration takes office.

To this complaint one could say that this is the way things are supposed to be in democratic societies. Citizens and elected representatives are allowed to change their mind about what values are important, and to want to have those changed views reliably reflected in what the government produces and how it behaves in its encounters with citizens. One could also observe that private sector executives also face fads and fickle demands for products and services to which they have to frequently adjust. So, once again, we may simply have to recognize that if we are to maintain the utility and relevance of performance measurement systems in meeting political demands for accountability, we will simply have to change them often.

But another answer is possible. What if it turned out that while the focus on particular dimensions of performance in the political world changed frequently, but that the changes were not really entirely new things, they were simply oscillations around a core set of values associated with a given organization. On this view, any given organization such as a police department is linked in the society's mind to a particular set of values that are judged to be at stake in the way the organization operates. We will always be interested in crime control, always be interested in protecting rights and calling offenders to account, always interested in reducing levels of fear, always interested in limiting the use of lethal force, always interested in building stronger relations with communities, always being responsive to citizens calls for service, etc. What happens in society, then, is not that wholly new values are added or subtracted from this mix, but that some come to be emphasized more than another at a particular point of time.

If this were true, one would expect that over some period of time – say a decade or so – the society would have shown its hand in nominating all the important values that were likely to be viewed as important over the next decade, and the next decade after that. Living through the decade, it might seem that the police were being pushed from pillar to post in terms of the values that they should seek to achieve and reflect in their operations. Yet, at the end of the decade, one would see that the demands for accountability were staying within a relatively stable core. Moreover, one might discover that even when people seemed to be concerned exclusively with crime control and indifferent to cost, or the use of force, or the protection of individual rights, they didn't really mean that those values should be abandoned. They just wanted them de-emphasized for a particular time. This would also mean that we wouldn't have to abandon the measurement systems developed to respond to urgent demands at an earlier time, we could sustain these sustains confident in the view that particularly urgent concern about this dimension of performance will come back again, that interest in this old dimensions of performance remains despite current priorities, and that if we sustain the measurement of the old values of performance even as we are measuring the new, we will have more information about the production possibility frontier in the future than we have now.[43]

If this claim about the stability of an underlying core set of values for given organizations proves accurate, then the idea that we might be able to develop a relatively stable set of measures that could reliably capture the public value produced by a given kind of public organization gains standing. All we would have to do is to keep adding dimensions of value to be measured until the political system ran out

[43] For a more sustained discussion of this idea, see Chapter 6 of Recognizing Public Value.

of questions to ask. We would then also have created an accurate and useful, if complex, view of what values the public writ broadly wanted to see achieved by and reflected in the operations of the organizations that were operating in their name, and using their collectively owned assets.

I have gone on at some length in developing ideas about the nature of the political work facing public managers in developing performance measurement systems and how certain kinds of problems can be managed or endured not because I am confident in the answers. We know much less about the intersection between politics, public demands for accountability, and performance measurement and management systems than we do about any of the other kinds of work. But I am pretty sure that the hopes for managing organizations to create public value cannot be realized unless we figure out how to bring the worlds of politics, external accountability, and performance measurement systems together.

C. Developing the Public Value Account

To begin this strategic work of developing performance measurement and management systems for government organizations that guide them towards increased public value, it is useful to set out a framework that can capture the most important issues to consider, and to accumulate important ideas and information as we move forward. One such framework is described in *Recognizing Public Value* as the Public Value Account. This framework is meant to do for government managers what the financial bottom line does for private sector managers: namely, provide a way of accounting for costs incurred and valuable results produced by a government organization. The difficulty, however, is that this scheme has to cope with all the following features of accounting for value in government:

First, it has to account for the fact that the assets that government uses to produce socially valuable results are not restricted to money. The government can alter material and social conditions in the world by using its authority as well as its money. It does this when it uses the draft to create an army, when it imposes clean-up burdens on polluting industries, when it seeks to reduce traffic fatalities by imposing speed limits. It also uses its authority to ensure that benefits made available by the society as a whole to specific classes of beneficiaries go to the intended beneficiaries and not to others. Such actions alter individual and social conditions in ways that produce benefits and losses to particular individuals, and that are or are not consistent with particular ideas of what we owe to one another as a matter of justice and prudence. Like money, authority is valued, and in short supply. All other things being equal, if the society can find a way to produce desired results using less authority, then that solution would be preferred to the one that made greater use of state authority.

Second, it has to recognize that individuals and their satisfaction are not necessarily the appropriate arbiter of public value. This is particularly true when the government is using its authority to oblige individual citizens to do something they would prefer not to do for the benefit of the community as a whole. But is also true when the government is providing benefits to individual clients, for example, when it provides welfare payments to poor women with children, or when it provides drug addiction services to addicts. In these cases, society as a whole may have the satisfaction of the individual clients in mind as an important objective. But that is usually not the only important objective. The society as a whole often wants something more than the happiness and satisfaction of the client: they hope the poor mother will provide adequate care for her children, and/or will find gainful employment, and they hope that addicts will give up their addiction, set a job, stop committing crimes, and care for their dependents.

Third, it has to recognize that when the government acts, it will be judged by deontological standards of fairness and justice as well as utilitarian standards of satisfying individual clients or achieving desired social outcomes. Individuals affected by government actions will want to be sure that their rights and interests have been appropriately protected from arbitrary action, and so will their fellow citizens who are concerned that the government act fairly and justly as well as expediently. All of these broad features. One of the main reasons for this fact is that these are the normative concerns that attach to the use of state authority, and it is nearly always true that when government is acting, it is acting using the significant authority of the state. This is obviously true when it is imposing obligations on individuals and enforcing rules about who is entitled to what particular government benefit. But is also true when it is simply spending money to produce a good or service that is available to all, because the money it is spending came into its hands through the use of authority – namely, the taxing power of the state. If concerns about justice and fairness attach themselves to any use of state authority, then presumably they attach themselves to the use of money that was raised through the use of state authority.

Working out the implications of these philosophical points for the construction of a "bottom line" for public agencies is a difficult but not insurmountable task. Let's assume at the outset that, as noted previously, government can have as good a cost accounting system as the private sector – at least with respect to the use of dollars to produce results and provide benefits. There should no problem in accounting for that and doing so in financial terms.

The problems begin, as noted above, when we try to find the equivalent of the revenue side of private sector accounting: the money earned by the enterprise through the sale of goods and services to willing clients with money to spend. That is generally missing in the public sector. If we wanted to take a customer perspective in government, we could, perhaps substitute some measure of satisfaction experienced by the clients of government agencies. But, as noted above, there are many clients of government agencies including criminals, polluters and tax cheats that receive obligations from the government rather than services, and whose satisfaction cannot be the only, and probably not even the principal aim of the encounter. And often, even when the government is providing services of some kind, it is at least as interested in achieving some desired aggregate social welfare that is not necessarily captured in the satisfaction of individual clients – whether obligatees or service beneficiaries.

There are also many individuals in particular positions in society who are not clients of particular government programs. Those individuals could include voters, or taxpayers or someone we might want to call a citizen in whose name the government was trying to act. We often have no sure way of knowing what values individual citizens would like to see achieved by and reflected in government operations.

What we rely on instead is a kind of collective utilitarianism in which we imagine that there is a public more or less appropriately constructed that can become articulate about what the public as a whole, or all of us acting as citizens rather than clients want to see happen in the operations of government agencies. We call this legislating or public policy-making. And it relies on their being both a collective entity that can express what the public values, and the likelihood that what the public values would consist not only of the satisfaction of individual clients interacting with the government, but in the achievement of desired social outcomes characterized by aggregate states of the society as well as individual satisfactions with their particular state. Put more simply, we give government enterprise policy mandates or missions that describe the important values they are to seek to produce and reflect in their operations.

Given this discussion, we can start with a public value account that records financial costs on the left hand side of the ledger. On the right hand of the ledger we can put the satisfaction of clients of government agencies as an important value created by the government. But the more important idea is the specific values that were described as the mission or the ultimate goals of the government enterprises. So far, so good. We have an idea of cost effectiveness in achieving a desired social goal, with some attention being given to the satisfaction of those who become clients of government enterprises. (Taxpayers are presumably mollified by the inclusion of a focus on cost, and citizens, we hope, are satisfied with the concern give to both cost and valued social result – just as corporate shareholders would be in interested in overall profit.)

The next step, however, is to recognize that authority is being used as an asset, and to begin paying attention to how much authority is being used in any particular government enterprise. It may seem odd to think of the use of authority and force as a quantitative idea, but on reflection, it seems less difficult than it might first appear. We can certainly think about the degree of force that is used to promote compliance with particular obligations. And we can think about the magnitude of the burden that is imposed by the enforcement of any given regulation. We can even turn these estimates into financial costs of enforcement and compliance. And on occasion, when the government is forced to pay compensation to those whose rights were violated, we can get a direct estimate of the financial of an improper use of force. This might give us a way to monetize other abuses of authority that do not reach this stage of public consideration. So, while it might strain our current practices to think about and try to account for the use of authority as an asset of government, it does not seem completely beyond the pale to start thinking about whether we could notice whether we were using more or less force to accomplish a particular result. Surely the citizens on whom the authority of the state is being used can notice degrees of force being deployed.

What follows closely on the observation that we have to account for the use of authority as a cost in government operations, is that once we recognize force is being used, we have to begin accounting for the fairness with which the government acts, and the degree to which its policies are pushing the society towards one or more particular images of a good and just society. This is necessarily true because in a democratic society, the use of state authority always has to justified. It cannot be used arbitrarily or unfairly. There has to be a reason, and the reason has to be a general rule that applies to all. Obviously, it is much easier to account for fairness in government access than its success in producing other important aspects of justice. But given that these matters of justice and fairness show up all the time in discussions about how government agencies are performing, and whether they are advancing justice in one form or other, it might be good to begin thinking about what facts could be gathered about its activities that pointed in one direction or another. To take the example of a police department again, one might want to develop measures of how much force and authority a given department was using to enforce the law, the extent to which it was successful in calling offenders to account for their crimes through arrest and successful prosecution, the fairness with which decisions to use unusually intrusive enforcement methods were made, the frequency with which it seemed to violate rights, and how fairly police resources seem to be allocated across both certain kinds of crime and geographic neighborhoods.

Taken together, the interests in financial costs to government, the achievement of collectively desired outcomes, the satisfaction of individual clients, the degree to which state authority was being used, and the degree to which government agencies were acting fairly and helping to achieve a just as well as a prosperous society would produce a public value account like the one pictured in Figure 3. In this version, I have added the paradoxical idea that a public organization might want to include room in the accounting scheme for recognizing unanticipated good and bad effects of its activities that were not part of its original mission. This is serve as a reminder that we can easily be surprised by what happens that matters to individuals when the government begins acting, and to create a category that can be used for adapting the public value account as we learn more about the consequences of the choices we have made.

[Insert Figure 3 About Here]

The challenge, of course, is to begin filling in this abstract frame with more particular ideas about public value – the dimensions of value that citizens and their representatives would like to see produced by and reflected in the operations of government agencies, and with concrete measures that could capture the degree to which these values were actually being produced in the world in ways that registered with citizens, voters, taxpayers and clients (both obligatees and beneficiaries). That will have to be done agency by agency over a period of time. Figure 4 presents one version of a Public Value Account for a the NYPD – an organization that has been on my mind for more than 30 years.

Perhaps over time, certain classes of agencies such as police departments, or schools, or social service agencies, may develop concepts of value and systems of measurement that are specific to their particular kind of work. That is, the scheme developed for the NYPD might be useable by other urban police departments, or maybe all police departments. That would be welcome development, since we could then begin to codify practices, improve measures, and make comparisons among different agencies in the same line of business in the interest of learning how to do better than we are now doing. Once data began being collected in similar systems capturing different dimensions of performance, We might actually be able to see the current production possibility frontier in particular public industries, which particular agencies are defining the frontier for particular industries, and what agencies have room to improve on all dimensions by moving out towards the frontier. But the start is to see if we can outline in sharper and sharper detail the target we are trying to hit: a reasonably satisfying conceptual definition of public value.

D. Deloping a Public Value Scorecard.

But the Public Value Account is just the first step. Just as private sector managers learned that their financial accounts alone were inadequate as a basis for evaluating and managing the performance of private sector organizations because they were too much about the past and not enough about the future, and too much about the ultimate ends and not enough about the processes that made the ultimate ends possible in the future, so public sector managers might learn that they need something more than the public value account to manage their enterprises. Shamelessly following the example of Kaplan and Norton, I proposed the creation of a Public Value Scorecard that included the public value Account (just as the Balanced Scorecard includes the Financial Perspective), but then add to that basic account families of measures that will focus managerial attention on the current position of the organization in its environment, whether and how operations can be sustained, and how both current operations and future positioning might be improved through particular investments that embodied the continuous learning of the organization. In constructing this family of measures, I relied on the strategic categories developed in *Creating Public Value*, and often presented as the "strategic triangle." (See Figure 4.) The Public Value Account (already discussed) captures the particular perspective associated with the "public value" circle of the "strategic triangle. " To this, I added perspectives associated with Maintaining and Building Legitimacy and Support (roughly analogous to Kaplan's customer perspective) and the Operational Capacity Perspective (roughly analogous to Kaplan's operations perspective).

[Insert Figure 4 About Here]

I did not create a separate category for the learning perspective since it seemed to me that this was the dynamic quality of all the perspectives. By thinking strategically and developmentally, we were taking it for granted that we were learning about what constituted the important dimensions of public value that should be accommodated in the Public Value Account, that we were learning both about our current standing with authorizers, and the legitimacy we enjoyed and how that might be bolstered, and that we were learning about how to improve our performance in current activities, and what new things might be added or old things subtracted to improve our overall performance.

Outlines for the Legitimacy and Support Perspective, and for the Operational Capacity perspective are presented in the abstract in Figures 5 and 6. Again, these abstract categories have to be given richer content both by developing more specific concepts, and measures attached to those different concepts. And, as we accumulate experience in constructing these sorts of measurement systems, some parts may become more easily standardized. For example, it is not hard to imagine that in the future many different public organizations would use simple account management systems to keep track of their relationship with key actors in their authorizing environment. To show more concretely what is meant by some of the abstract categories, Figures 8 and 9 present ideas about how these different perspectives might be filled in by the NYPD.

III. Conclusion

Citizens in democratic societies, and their elected representatives have long sought the simplicity of the private sector's bottom line as a basis for improving the performance of government.

And they have often assumed that such a thing was easily within reach, and it was only the resistance of self-protective bureaucrats that prevented the creation of bottom line management systems for government.

I don't think citizens are wrong to want good measurement systems that can enable effective public oversight and accountability of public organizations. Indeed, I count myself among the most ardent enthusiasts of both political and governmental accountability to citizens in democratic societies. And I have devoted a great deal of observation and thought to imagining how such a thing could be accomplished.

Where I end up parting company with the most enthusiastic advocates of bottom line management for government is not in the desire to achieve that goal, but in the assumption that such a feat is simple and immediately within reach. Oliver Wendell Holmes once said, "I wouldn't give a fig for the simplicity that lies on this side of complexity. But I would give my right arm for the simplicity that lies on the other side of complexity." I think, like Holmes, that we ought to hold out for the simplicity that lies on the other side of complexity in seeking accurate and useful measures of public value creation. To get there, we have to fight through not only the technical issues of how to construct reliable empirical measures of concepts we have named as valuable, and the managerial issues about how to use such measures in motivating performance in organizations and learning how to improve. We also have to take up and answer the tough political and philosophical questions that lie behind any judgment of value in a democratic regime. We have to think about how a collective entity that could reasonably described as a public can be called into existence and become articulate about what it would like to achieve not only in the large, but also in the small through governmental action. And we have to keep talking about the important issues of justice – about the kinds of relationships we would like to have exist in our society, and what as a consequence we might owe to one another as something that government might help us achieve. The challenge, it seems to me, is to find ways to have that rich, swirling discussion touch down periodically on the ground through the medium of measurement systems that seek to capture a current understanding of what public value we seek to produce.

Figure 1:

Accountability Agents and Dimensions of Value: NYPD

Formal Oversight: Continuous						
Federal Government	—	—	—	—	—	—
Governor	—	—	●	●	●	●
State Legislature	—	—	●	●	●	●
Mayor	●	—	●	●	✱	●
Office of Operations	—	—	●	●	—	—
Office of Management and Budget	—	—	—	○	—	—
Criminal Justice Coordinator	—	—	○	—	—	—
Law Department/Carpenter Council	—	—	—	—	○	○
Office of Investigation	—	—	—	—	○	○
City Council	●	—	—	●	—	●
Community Boards	○	—	○	—	—	—
Civilian Complaint Review Board	—	—	—	—	●	—
Police Commission	—	●	●	●	●	●
Formal Oversight: Specially Commissioned						
Mollen Commission	—	—	—	—	✱	—
Gighenti Crown Heights Commission	✱	✱	—	—	✱	✱
Informal Oversight: Citizens and Nonelected Representatives						
Electronic Media	—	✱	✱	—	—	—
Print Media	—	✱	✱	—	—	—
Citizens Budget Commission	—	—	—	●	—	—
ACLU	—	—	—	—	●	—
Police Unions	—	—	—	—	—	—
Alliance for Professional Policy	—	—	—	—	○	—
Community Groups	○	—	—	—	—	—
Informal Oversight: Partners in Criminal Justice						
Courts	—	●	—	—	—	—
District Attorneys	—	●	—	—	—	—

● – *Continual Intense Oversight* ○ – *Continual Less Intense Oversight*
✱ – *Intermittent Intense Oversight* ◌ – *Intermittent Less Intense Oversight*

Figure 2:

The Value Chain

Possible Points of Measurement

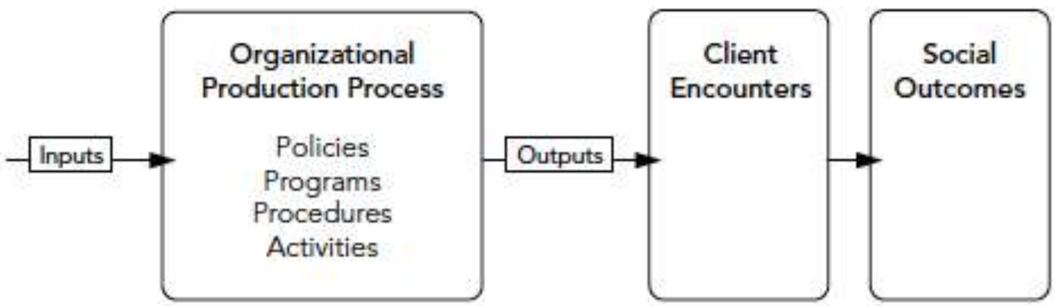


Figure 3:

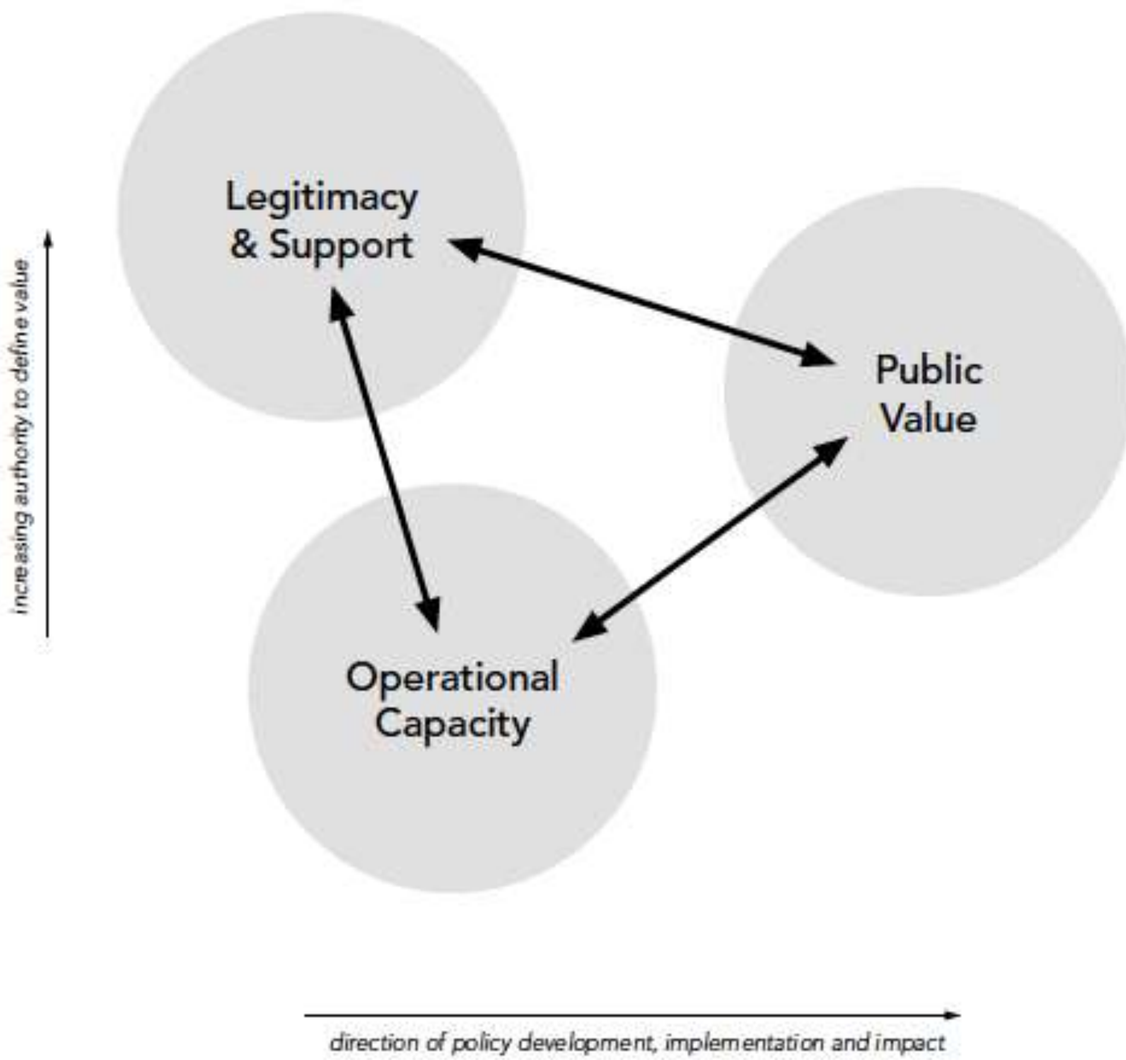
A Schematic Version of a Public Value Account

PUBLIC VALUE ACCOUNT	
Use of Collectively Owned Assets	Achievement of Collectively Valued
and Associated Costs	Social Outcomes
Financial Costs	Mission Achievement
Unintended Negative Consequences	Unintended Positive Consequences
	Client Satisfaction
	Service Recipients
	Obligatees
Social Costs of Using State Authority	Justice and Fairness
	At Individual Level in Operations
	At Aggregate Level in Results

Figure 4:

The Strategic Triangle

for Imagining and Testing Public Value Propositions



THE LEGITIMACY AND SUPPORT PERSPECTIVE: General Form

Mission Alignment with Values Articulated by Citizens
(Link to Public Value Account)

Inclusion of Neglected Values with Latent Constituencies
(Link to Public Value Account)

Standing with Formal Authorizers:
Elected Executives
Statutory Overseers in Executive Branch (Budget, Finance, Personnel)
Elected Legislators
Statutory Overseers in Legislative Branch (Audit, Inspectors-General)
Other Levels of Government
Courts

Standing with Key Interest Groups:
Economically Motivated Suppliers
Self-Interested Client Groups
Policy Advocacy Groups
Latent Interest Groups

Media Coverage:
Print
Electronic
Social

Standing with Individuals in Polity:
General Citizenry
Taxpayers
Clients
- Service recipients
- Obligatees

Position of Enterprise in Democratic Political Discourse:
Standing in Political Campaigns
Standing in Political Agendas of Current Elected Regime
Standing in Relevant "Policy Community"

Status of Key Legislative and Public Policy Proposals to Support Enterprise *(Link to Operational Capacity Perspective)***:**
Authorizations
Appropriations

Engagement of Citizens as Co-Producers *(Link to Operational Capacity Perspective)*

THE OPERATIONAL CAPACITY PERSPECTIVE: General Form

Flow of Resources to Enterprise *(Link to Legitimacy and Support Perspective)*:

Financial Revenues Flowing to Public Agencies:

- Appropriations
- Intergovernmental grants
- Fees

Legal and Statutory Authorizations/Mandates

Public Support/Popular Opinion

Human Resources:

Current Status of Workforce

- Size
- Quality
- Morale

Recruitment and Selection Processes

Training/Professional Development of Staff

Compensation Levels

Advancement Opportunities

Performance Measurement Systems for Individual Accountability

Public Volunteer Efforts

Operational Policies, Programs and Procedures:

Quality of Operational Performance

- Documentation of current procedures
- Compliance with tested procedures
- Auditability of performance recording methods

Organizational Learning

- Evaluation of current untested policies
- Stimulation and testing of innovations
- Institutionalization of successful innovations

Internal Resource Allocation

Performance Measurement and Management Systems

- Investment in systems
- Use of systems

Organizational Outputs *(Link to Public Value Account)*

Quantity of Outputs

Quality of Outputs

- Attributes that produce desired results
- Attributes that increase client satisfaction
- Attributes that reflect justice and fairness in operations