

MENSURAÇÃO DE DESEMPENHO NO SETOR PÚBLICO: OS TERMOS DO DEBATE*

*Regina Silvia Pacheco***

INTRODUÇÃO

DUAS ORIENTAÇÕES DISTINTAS

POLÊMICAS E PROBLEMAS FREQUENTES

**“DOENÇAS” DA MENSURAÇÃO DE RESULTADOS NO
SETOR PÚBLICO**

LIÇÕES APRENDIDAS

Referências

* Este trabalho está publicado na Biblioteca Virtual TOP em permissão do autor.

** Professora da EAESP-FGV - Escola de Administração de Empresas de São Paulo, da Fundação Getúlio Vargas e coordenadora do Mestrado Profissional em Gestão e Políticas Públicas. Foi presidente da ENAP Escola Nacional de Administração Pública entre 1995 e 2002. Contato: regina.pacheco@fgv.br

MENSURAÇÃO DE DESEMPENHO NO SETOR PÚBLICO: OS TERMOS DO DEBATE

Regina Silvia Pacheco

INTRODUÇÃO

O objetivo deste trabalho é contribuir para a sistematização do debate sobre mensuração de desempenho no setor público. O tema é polêmico e sua implementação nas duas últimas décadas tem sido acompanhada de inúmeros artigos e estudos.

Para alguns autores, a mensuração de desempenho no setor público tem contribuído para o alcance de múltiplos objetivos, dentre eles a transparência de custos e de resultados, a melhoria da qualidade dos serviços prestados, a motivação dos funcionários, sendo um dos pilares mais importantes da nova governança em torno do Estado-rede (Goldsmith & Eggers, 2006; Behn, 1995). Para seus críticos, no entanto, trata-se de uma transposição indevida de um instrumento desenvolvido para a gestão empresarial, que gera graves distorções quando aplicado ao setor público (Dunleavy & Hood, 1994; Hood, 2007).

Apesar da polêmica, as experiências de mensuração de desempenho e contratualização de resultados têm-se expandido tanto em outros países como no Brasil; diferentes governos vêm enfrentando as dificuldades introduzidas por esta nova forma de gestão e procurando aperfeiçoar metas e indicadores.

A mensuração de desempenho é central nas mudanças em curso. É o pilar de sustentação de duas das três inovações responsáveis pela melhoria de desempenho das organizações públicas, segundo pesquisa comparada compreendendo sete países: a contratualização de resultados e o orçamento por produto aliado à contabilidade gerencial (Jann & Reichard, 2002) ¹

Mensurar resultados é uma ferramenta que vem sendo parte de um conjunto abrangente de mudanças que incluem a revisão da macro-estrutura do Estado e a criação de “*arm’s lenght organizations*” (agências e organizações públicas não-estatais), a definição prévia de resultados a alcançar, a concessão de flexibilidades à organização que se compromete previamente com resultados, o reconhecimento do papel do *public manager* a quem é concedida maior autonomia e imputada nova responsabilização pelos resultados visados; tais mudanças são acompanhadas portanto de uma nova distribuição de responsabilidades, ou *accountability* por resultados.

A montante, tais inovações requerem o recurso ao planejamento estratégico a fim de que sejam clareados os objetivos e fixados os resultados visados; a juzante, novas formas de controle são desenvolvidas, menos baseadas no controle formal de procedimentos, e mais voltadas à comparação de resultados obtidos por organizações similares, ou simplesmente maior transparência quanto ao uso dos recursos públicos por meio do acompanhamento dos resultados alcançados. Esta é também a base para o desenvolvimento de novas formas de relacionamento entre entidades públicas e parceiros – públicos, públicos não-estatais ou privados.

¹ Os países pesquisados têm diferentes histórias administrativas e contextos políticos: Dinamarca, EUA, Holanda, Nova Zelândia, Reino Unido, Suécia e Suíça. A terceira inovação identificada pela pesquisa refere-se a mudanças nas relações de trabalho, diminuindo as diferenças entre os contratos de trabalho nos setores público e privado e introduzindo flexibilidades na gestão de pessoas.

DUAS ORIENTAÇÕES DISTINTAS

Apesar de parte da literatura remeter a gestão pública por resultados a uma única visão – a que vê a gestão privada como superior e quer introduzir seus métodos no setor público, há autores que identificam pelo menos duas visões distintas como fontes de inspiração das mudanças em curso.

Kettl (1997) organiza essas duas correntes em torno de dois lemas distintos: por um lado, “*make managers manage*”, reunindo países que criaram incentivos visando influenciar comportamentos; por outro, “*let managers manage*”, expressando a visão de que há inúmeras barreiras a serem removidas, regras, procedimentos e estruturas rígidas que impedem o administrador público de administrar. No primeiro caso, a contratualização de resultados representa uma nova forma de controle, e vem acompanhada do estabelecimento de sanções positivas e negativas; o país que levou mais longe tal perspectiva é a Nova Zelândia, com a primeira geração de reformadores. No segundo caso, o acordo de resultados é visto como instrumento de coordenação, ajuste e aprendizado organizacional; a experimentação, e não o controle, é a aposta para a melhoria do desempenho; esta tem sido a marca das reformas na Austrália e Suécia.

De maneira similar, Jann e Reichard (2002) também fazem referência a dois grupos de inspiração distinta, ao analisar a contratualização de resultados. Os autores identificam um dos grupos como os defensores da eficiência (ou minimizadores do custo), o outro como reformadores em busca da melhoria de desempenho do setor público (maximizadores dos resultados); o primeiro grupo recorre a mecanismos de punição e recompensas de acordo com o desempenho alcançado; para o segundo, o alvo é o aprendizado baseado em relações de confiança como caminho para alcançar melhores resultados. Os autores identificam a escolha racional e a teoria da agência como referência do primeiro grupo, ao adotar o contrato de resultados como um novo instrumento de controle para enfrentar o problema do comportamento maximizador do auto-interesse; na segunda vertente, é destacado o comportamento cívico dos agentes, que dá sustentação à lógica do aprendizado mútuo.

Trosa (2001) também destaca diferentes visões acerca da mensuração de resultados e seu uso por diferentes países e governos. Suas análises são análogas às já citadas; no caso do uso radical da mensuração e contratualização de resultados como nova forma de controle, a autora vê o risco de um novo formalismo e rigidez. Após ter acompanhado a experiência de vários países, Trosa conclui que os resultados são melhores quando as metas são negociadas (e não impostas unilateralmente) e quando as flexibilidades são concedidas tendo metas a alcançar como contrapartida.

As orientações sobre o tema são tão distintas que levaram alguns autores a defender a existência de uma forma de “mensuração de desempenho à la nórdicos”². Para eles, a perspectiva nórdica afasta-se da anglo-saxã, já que esta associa desempenho a incentivos financeiros individuais, enquanto que aquela valoriza o aprendizado:

“Given the egalitarian values, many Scandinavians may well be puzzled by the importance imputed to individual financial incentives in the Anglo-American culture. Organisational learning ... may better suit the Nordic approach. Naturally, in the Nordic countries too, individuals are assessed and rewarded for their contribution to organisational effectiveness, viability and so on, but probably not as extensively, and certainly not as

² uma edição especial da revista *Financial Accountability and Management* foi dedicada ao tema [22(3), 2006], introduzida por Johnsen, Norreklit e Vakkuri (2006), que explicitam e defendem “a nordic perspective on public sector performance measurement”.

taken-for-granted in such a manner as seems to prevail in some Anglo-American cultures...” (Johnsen, Norreklit e Vakkuri, 2006, p.208).

Distinguir estes dois grupos de inspiração pode ajudar a clarear o debate no Brasil. Aqui, muitos críticos referem-se apenas às experiências mais difundidas, dos países anglo-saxões, que levam a atribuição de recompensas e penalidades até o nível individual. Também é interessante que os formuladores da política de gestão conheçam a experiência dos países nórdicos e possam optar por um modelo ou outro, conhecendo as implicações de ambos.

POLÊMICAS E PROBLEMAS FREQUENTES

Apesar de vários autores ressaltarem a importância da gestão por resultados para a melhoria do desempenho do setor público, apontam também polêmicas e problemas frequentes associados a essa nova forma de gestão. Tais problemas não têm justificado o abandono do modelo; têm, ao contrário, levado ao seu aprofundamento, buscando corrigir rotas e superar os obstáculos identificados.

Parte dos problemas relacionados à mensuração de desempenho pode ser atribuída à adoção prematura ou isolada apenas desta ferramenta, desvinculada do conjunto de mudanças que a acompanham e que constituem elementos de uma ampla reforma da organização e funcionamento do Estado. Assim, são grandes as chances de insucesso quando a mensuração de desempenho é adotada sem que as demais inovações na gestão estejam presentes. O grande risco é que a mensuração se torne um fim em si mesmo, desvinculada do objetivo maior que é a melhoria do serviço público prestado ao cidadão (Behn, 1995).

Além deste aspecto, a adoção parcial e isolada da mensuração de desempenho, é possível sintetizar as polêmicas em torno da mensuração de resultados em três grupos de temas: a discussão em torno do que mensurar – produtos (*outputs*) ou impactos (*outcomes*); a adoção de sanções positivas e negativas; a vinculação de parte da remuneração individual ao desempenho.

Em texto anterior, sistematizamos o debate em torno do que mensurar – *outputs* x *outcomes* (Pacheco, 2006). A defesa da mensuração de impactos (*outcomes*) ou a contribuição efetiva para a resolução de um problema, tem levado alguns analistas, tanto no Brasil como em outros países, a criticar boa parte das experiências em curso, já que a maioria delas se inicia pela mensuração da prestação de determinados serviços previamente especificados (*outputs*). Mas há autores que consideram tal debate inócuo; para esses autores, ambas as opções apresentam vantagens e inconvenientes. Na mensuração de impactos é difícil estabelecer relações de causalidade entre as ações empreendidas e o resultado observado – é difícil isolar, dentre as inúmeras variáveis que afetam a situação, aquelas diretamente ligadas aos serviços prestados por uma determinada organização pública. Por vezes as ações requerem longo tempo de maturação para que os impactos possam ser observáveis. Os impactos desejados podem depender de mudanças substanciais no comportamento dos cidadãos. Por outro lado, mensurar produtos pode levar a um foco excessivo no curto prazo.

Para Trosa (2001), ambas formas de contratação são importantes e respondem a duas perguntas distintas. A mensuração de *outputs* permite conhecer o que é efetivamente produzido com os recursos públicos; já a preocupação com *outcomes* ou impactos permite indagar sobre a eficácia e utilidade daquilo que é produzido. Segundo a autora, a resposta a esse debate deve ser pragmática: governos devem começar pelos serviços prestados, cuja mensuração é mais fácil, e ir evoluindo em direção aos impactos, por meio da construção da

cadeia lógica que liga as ações aos objetivos visados, relacionando impactos, resultados intermediários e ações.

Já Behn (2004) defende enfaticamente a mensuração de outputs, dentre outras razões porque é um instrumento poderoso a ser utilizado pelo dirigente de uma organização pública para motivar seus funcionários; para este autor, o objetivo do administrador é motivar, enquanto que o objetivo do economista é controlar. Outra vantagem da opção pela mensuração de produtos é que deste modo é possível definir metas claras a serem buscadas; o autor ressalta que cabe à atividade de avaliação do programa estabelecer os elos entre os impactos visados e os produtos ou serviços a serem prestados, de forma que os nexos lógicos entre estes e aqueles sejam sempre explicitados.

Assim, não há evidências que comprovem a superioridade da mensuração de impactos sobre a mensuração de produtos. As tentativas de opor as duas formas de medidas parecem não fazer sentido, à luz das lições aprendidas da experiência internacional. Países que contrataram extensivamente resultados via produtos (outputs), como Nova Zelândia, trataram de corrigir excessos introduzindo a mensuração de impactos; países que privilegiaram a mensuração de outcomes, como Austrália, deixaram espaço amplo demais às organizações sem cobrar delas compromisso com ações diretamente mensuráveis.

Também há variações de um setor a outro. Na saúde, desenvolveram-se os indicadores de produto e de resultados intermediários, e as cadeias lógicas entre produtos e impactos estão mais claramente estabelecidas e aceitas. Na educação, tem-se multiplicado as políticas onde a mensuração de impactos é adotada via avaliação externa do rendimento dos alunos.

No Brasil, o setor de saúde parece muito mais preparado para conviver com a mensuração do que a área da educação, talvez porque a própria lógica de remuneração do SUS, via procedimentos, tenha aberto caminhos para a mensuração de serviços prestados. Segundo depoimento de um gestor municipal, o foco em procedimentos foi tão disseminado pelo SUS que hoje impede avanços para uma visão de “linha de cuidados” ou para procedimentos não finalísticos como supervisão médica ³. Na Grã-Bretanha, a política de mensuração de resultados em saúde se aprofunda, com a adoção recente da mensuração da satisfação do paciente sobre a qualidade da atenção médica recebida, com impactos para a remuneração do médico.

Já na área da educação, no Brasil, a adoção de avaliação externa da aprendizagem dos alunos (impacto) tem provocado reação aguda dos sindicatos de professores, com exceção talvez de Minas Gerais⁴. A principal alegação de lideranças sindicais e de especialistas em pedagogia diz respeito à autonomia do professor, que seria desrespeitada ao impor a ele metas de aprendizagem a serem alcançadas por seus alunos. Também se alega impossibilidade de atribuição de causalidade entre a atuação do professor e o rendimento do aluno, que estaria condicionado por múltiplos fatores extra-classe, tornando impossível medir o impacto da ação do professor – em sintonia direta com o que foi relatado antes sobre as dificuldades de mensuração de *outcomes*.

Médicos e professores da rede pública fazem parte do segmento denominado por Lipsky (1980) como *street level bureaucrats*, ou funcionários de ponta, especializados, para quem a deontologia de sua profissão está acima de seu vínculo com o Estado, e que resistem a serem submetidos a controles externos. A adoção de medidas de desempenho provoca

³ Palestra da Secretária-adjunta da Saúde de Curitiba, GV Saúde, abril de 2009.

⁴ A maior aceitação, pelos professores mineiros, pode ter como parte da explicação o fato de se tratar de uma política abrangente de governo para todos os setores; em outros casos, trata-se de uma política setorial isolada.

sentimento de perda de autonomia, que tende a mobilizar reação contrária à política. No entanto, temos visto maiores avanços na saúde do que na educação.

Uma outra explicação para esta diferença, além da já assinalada, pode estar associada ao status do gerenciamento nos dois setores: enquanto o diretor de hospital é já um papel aceito e destacado entre profissionais da área da saúde, o diretor de escola pública é um entre pares, não sendo visto como autoridade legítima responsável pelos resultados da escola que dirige e portanto sem poder de direção sobre os professores que ali atuam. Outro motivo pode estar vinculado à introdução do novo formato organizacional na saúde – a mensuração de desempenho difunde-se no momento em que são criados os hospitais no novo formato de organização social de saúde, enquanto que as escolas públicas permanecem como parte da administração direta onde ainda impera a ausência praticamente total de autonomia de gestão.

Na área da cultura a mensuração de resultados está apenas começando, também ligada à adoção do formato “organização social”. No estado de São Paulo, adotou-se indicadores de produção, bastante simplificados e homogêneos para instituições de natureza distinta. Assim, o desempenho de museus passou a ser medido em termos de nº de dias abertos ao público, nº de exposições realizadas, nº de visitantes, sem distinção entre um museu de vocação para grande público e outro especializado. Um equipamento cultural dedicado à interação entre arte e tecnologia não conseguiu fazer valer o nº de acessos ao seu *site* como indicador de presença junto ao público – insistindo a Secretaria Estadual da Cultura em contar apenas as visitas físicas ao local.

O debate sobre o que e como medir não chegou à imprensa, a não ser por meio do conflito entre lideranças sindicais dos professores e secretaria da educação. Assim, a polêmica fica restrita aos especialistas e afetados pela mensuração, sem que a opinião pública ou os beneficiários do serviço público possam se envolver e pressionar por avanços. Em alguns casos, quando usuários tomam conhecimento do sistema de metas, passam a pressionar por punições em caso de não atingimento das mesmas – o que nem sempre é o fim visado pelo gestor da política de mensuração de resultados.

O ex-secretário de Modernização Administrativa de Santo André relatou as dificuldades que enfrentou junto aos representantes dos usuários ao adotar a mensuração de resultados nas unidades básicas de saúde – queriam punição aos funcionários a cada vez que uma meta não era cumprida; no entanto, o objetivo da política municipal adotada era o de promover o aprendizado sobre como melhorar o serviço prestado, e não o de aplicar punições em caso de dificuldades com o cumprimento das metas ⁵.

Na seção anterior, apresentamos as duas visões sobre o tema das recompensas e punições. Tais visões se desdobram também na questão da remuneração variável. Assim, os defensores da “perspectiva nórdica” consideram que os problemas da mensuração de desempenho aparecem quando se vincula resultados a recompensas financeiras:

“In fact, much of the perverse learning from performance measurement may be due to using performance measurement in close conjunction with organisational objectives and financial rewards. Performance measurement has a number of functions, including enhancing transparency, improving organisational learning and appraising performance. The more performance measurement is used compulsively for more of these functions, the more the performance measurement will be experienced as unfair ” (Johnsen, Norreklit e Vakkuri, 2006, p.208).

⁵ Marcio Bellisomi, palestra, EAESP-FGV.

A associação entre mensuração de desempenho e recompensas financeiras é menos difundida do que se crê: segundo pesquisa realizada pela OCDE junto a seus países-membros⁶, apenas 18% ligam orçamento a resultados (Nova Zelândia, Holanda...), poucos têm mecanismos formais de premiação ou punição segundo resultados; 41% não têm; só 11% dos países sempre vinculam salário ao atingimento de metas e 26% fazem tal vinculação esporadicamente⁷.

Outro estudo ressalta que o sucesso de sistemas de remuneração variável segundo desempenho depende de fatores como transparência na definição de metas, nas regras segundo as quais os funcionários serão avaliados e nas relações entre medidas de desempenho e remuneração (Perry et al., 2009, p.14).

Portanto, mensurar desempenho não implica sempre adotar remuneração variável ou recompensas financeiras segundo o desempenho. O tema é polêmico e imerso em controvérsias. Há um número significativo de países que adotam a mensuração de desempenho e divulgam os resultados obtidos, globalmente pelo governo ou setorialmente; apenas parte destes países têm atribuído recompensas financeiras a funcionários e dentre os que o fazem, muitos adotam valores de impacto moderado sobre a remuneração, a fim de não provocar desvios de comportamento importantes.

Em outros casos, aprofunda-se as consequências a partir da avaliação de desempenho. O presidente Obama, em seu primeiro pronunciamento sobre a política educacional nos EUA, afirmou que o mau desempenho de estudantes em avaliações externas de aprendizagem poderá levar à demissão de professores – dando seguimento e aprofundando a política de seu antecessor denominada “No Child Left Behind” (NCLB), que introduziu metas para aprendizagem de estudantes em escolas públicas em todo o país.

No Brasil, um dos problemas relacionados à remuneração variável por desempenho tem sido sua adoção isolada de outras medidas reformadoras. Na administração federal, a remuneração variável por desempenho foi generalizada no segundo governo FHC, sem que tenham avançado as outras medidas propostas de contratualização de resultados; no governo Lula, os valores variáveis foram aumentados significativamente, passando a representar mais de 50% da remuneração total em muitos casos, para em seguida reverter tudo em aumento salarial incorporando 100% da parcela variável aos salários (transformados em “subsídio”).

“DOENÇAS” DA MENSURAÇÃO DE RESULTADOS NO SETOR PÚBLICO

Além das polêmicas frequentes em torno da mensuração de resultados, a literatura aponta problemas inerentes às formas de mensuração de resultados no setor público. Assim, “a plethora of unintended, negative consequences are documented. Teachers are teaching the test; ambulances are waiting outside of the hospital to improve response times; waiting lists are reduced by creating waiting lists for waiting lists; follow up visits are cancelled; delayed trains are wrongly registered as broken and ‘creative accounting’ is abound ...” (Van Dooren, 2008).

Bouckaert e Balk (1991) sistematizaram um conjunto de problemas relativos à medida de desempenho em organizações públicas, apresentando-os como “doenças” e suas prováveis “curas”. Os autores organizam as doenças com relação a três diferentes aspectos: 1) alegações para não medir o desempenho; 2) problemas com a percepção das medidas; 3) problemas relacionados às próprias medidas. Ao todo, os autores apontam doze tipos de “doenças” relacionadas à mensuração de desempenho no setor público.

⁶ Dos 30 países-membros da OCDE, 27 responderam à pesquisa.

⁷ OECD / World Bank Budget Practices and Procedures Database, 2003.

As justificativas para não medir abrangem a “doença do Dr. Pangloss”⁸, ou a insistência em não medir algo por acreditar que aquilo não existe; “doença da impossibilidade”, que alega ser impossível medir; “hipocondria”, que considera que no setor público não se deve vangloriar de algo que tenha resultados positivos.

As doenças relativas à percepção dos números e volumes abrangem “doença do côncavo / convexo”, que leva à percepção aumentada ou diminuída do que está sendo medido; “hipertrofia / atrofia”, na qual o ato de medir estimula a produção desnecessária de mais *output* ou sua redução indesejada; e a “doença de Mandelbrot”, que ignora o fato de que o resultado da medida depende de como se mede.

Quanto às medidas propriamente, os autores elencam as seguintes “doenças”: “poluição”, ou o fato de misturar diferentes elementos daquilo que está sendo medido; “inflação”, ou o uso desnecessário de grande número de medidas; “doença dos iluminados ou *top-down*”, pela qual os dirigentes decidem sozinhos o que e como medir e impõem o conjunto à organização sem suficiente comunicação, levando à desmotivação dos funcionários; “doença do curto-prazo”; “miragem”, quando se mede algo diferente do que se considera estar medindo; e a “doença de desvio de comportamento” (*shifting disease*), causada por medidas que não contemplam a finalidade da organização e acabam provocando comportamento adverso.

Parte desses problemas é também ressaltada por outros autores. Goldsmith e Eggers (2006) referem-se a competição entre diferentes provedores pelos casos mais fáceis, visando melhorar seu desempenho medido. Ou ainda a prejuízos à qualidade do serviço quando as medidas enfatizam economias, o que pode levar a cortes de custo que prejudicam a qualidade. Os autores também fazem referência à polêmica sobre o que medir – insumos, produtos ou impactos. E apontam problemas decorrentes da ausência de dados confiáveis sobre a situação inicial a ser medida – quando sub-avaliados, podem levar a metas frouxas, muito fáceis de atingir, que podem provocar críticas da imprensa, por exemplo.

Outros autores referem-se ao *cream skimming*, a atitude de prestar os serviços mais fáceis ou atender os usuários mais fáceis, de modo a obter tempos de atendimento menores ou maior número de casos atendidos.

Jann e Reichard (2001) também relataram problemas referentes a metas pouco ambiciosas e falhas no seu monitoramento, dentre outros problemas frequentes.

LIÇÕES APRENDIDAS

Em boa parte dos problemas identificados na literatura, mesclam-se as dificuldades de mensuração de resultados no setor público com a introdução de formas de contratualização de resultados, pilar das reformas em curso. Ainda assim, muitos autores consideram ambas como avanços irreversíveis, que continuarão na agenda da modernização do Estado nos próximos anos (Lapsley, 2008).

Como buscamos apontar, as dificuldades relativas à mensuração de desempenho no setor público são substantivas, tanto no Brasil como nas experiências internacionais em curso. Não se trata de atitudes intencionais voltadas à manipulação de dados ou falseamento dos resultados; são antes dificuldades ligadas à natureza das atividades desempenhadas pelo Estado e ao fato de estarmos diante de um grande empreendimento reformador, do qual a mensuração de resultados é apenas um dos instrumentos – isolá-la do conjunto das reformas em curso é uma distorção que pode comprometer seus resultados.

Os comportamentos adversos não derivam, como vimos, de intenções individuais; são explicados pela assimetria de informações ou por outros construtos da teoria dos jogos,

⁸ Em referência à obra de Voltaire, *Candido*.

podendo estar presentes em diferentes contextos institucionais ou tradições político-administrativas – não são portanto derivados de práticas clientelistas ou desvios de conduta. Na prática, no Brasil, nenhuma denúncia de grande vulto foi associada à mensuração de desempenho ou aos novos formatos organizacionais que a vêm adotando (organizações sociais).

Apesar das dificuldades, o sentido das experiências em curso é o de aprofundar as reformas na direção da gestão voltada para resultados. Isto implica enfrentar as dificuldades de mensuração de desempenho, ultrapassando as justificativas simplistas para não medi-lo ou as críticas superficiais que vêm nestas iniciativas a “privatização do Estado”, simplesmente porque emprestam técnicas antes aplicadas pelas empresas privadas. Adaptações são requeridas, assim como esforços continuados de aperfeiçoamento das medidas e dos medidores, já que há evidências suficientes que comprovam a contribuição da mensuração de desempenho para a efetiva melhoria de resultados alcançados pelas organizações públicas – principalmente aquelas que prestam serviços, onde é mais fácil mensurar atividades e produtos do que nas secretarias formuladoras de políticas públicas.

As contribuições da mensuração de resultados para a transparência são notáveis. Por mais problemas que tenham os indicadores, é melhor contar com eles do que não tê-los. A emenda à lei orgânica do Município de São Paulo, proposta por iniciativa de um conjunto de entidades da sociedade civil, obriga os prefeitos eleitos a partir de 2008 a traduzirem suas promessas de campanha em um plano com metas e indicadores de desempenho. O conjunto de 223 metas apresentadas à cidade no último dia de março constitui uma base para o acompanhamento das ações de governo pela sociedade ⁹. É possível que haja várias das doenças mencionadas – inflação, poluição, *top-down*; não há metas de impacto; algumas são excessivamente genéricas, outras pontuais demais – ainda assim, melhor tê-las e debê-las em audiências públicas em todas as regiões da cidade do que não tê-las.

O mesmo pode ser dito dos avanços recentes da Previdência Social. Ter indicadores e metas sobre o tempo de concessão de benefícios é requisito indispensável para chegar a conceder aposentadorias em trinta minutos. Pode haver *cream skimming* ou perda de qualidade, caso não haja outros indicadores associados ao tempo de concessão do benefício. Mas uma cesta de indicadores combinados pode contribuir para evitar a *shifting disease*.

Outras experiências em curso merecem atenção por serem abrangentes e constituírem elemento central de uma política pública voltada para a gestão – os casos das prefeituras de Curitiba e Porto Alegre e o governo de Minas Gerais. São laboratórios a serem acompanhados, de onde sairão várias lições sobre possibilidades de melhorar o desempenho do setor público, trabalhando e retrabalhando medidas de desempenho, tanto de *outputs* como de *outcomes*, como um dos elementos de um conjunto de ações de modernização do Estado.

Resumo

A mensuração de resultados constitui um dos pilares das reformas em curso em diversos países nas últimas duas ou três décadas, sendo a base para outras inovações como agencificação, contratualização de resultados, remuneração variável por desempenho, parcerias, Estado-rede e outras. As polêmicas em torno da aplicação desta ferramenta de gestão no setor público são intensas – tanto entre adeptos como entre críticos e defensores. Este texto busca sistematizar os termos do debate sobre mensuração de desempenho no setor público, a partir da literatura internacional e das visões presentes no debate brasileiro. O objetivo é contribuir para o avanço do debate.

⁹ Ver www.prefeitura.sp.gov.br/agenda2012

REFERÊNCIAS

- Behn, Robert D. (1995). "The Big Questions of Public Management". *Public Administration Review*, 55(4):313-24.
- Behn, Robert D. (2004). "Why Public Managers Must Measure Outputs". *Public Management Report*, vol.1, nº10, June.
- Dunleavy, Patrick, and Christopher Hood (1994). From Old Public Administration to New Public Management. *Public Money and Management* 14(3): 9-16.
- Goldsmith, Stephen e William Eggers (2006). *Governar em Rede. O novo formato do setor público*. Brasília: ENAP.
- Hood, Christopher. (2007). "Public Service Management by Numbers: Why Does it Vary? Where Has it Come From? What Are the Gaps and the Puzzles?" *Public Money and Management*, 27: 95-102.
- Jann, Werner e Christoph Reichard (2002). "Melhores práticas na modernização do Estado". *Revista do Serviço Público* 53(3):31-50.
- Johnsen, Age, Hanne Nørreklit e Jarmo Vakkuri (2006). "Introducing a Nordic Perspective on Public Sector Performance Measurement". *Financial Accountability and Management* 22(3):207-12.
- Kettl, Donald F. (1997). "The Global Revolution in Public Management: Driving Themes, Missing Links". *Journal of Policy Analysis and Management*, 16(3): 446-62.
- Lapsley, Irvine (2008). "The NPM Agenda: Back to the Future". *Financial Accountability and Management* 24(1):77-96, February.
- Lipsky, Michael (1980). *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. New York: Russell Sage Foundation.
- Pacheco, Regina Silvia (2006). "Brasil: avanços da contratualização de resultados no setor público". Texto apresentado ao *XI Congresso Internacional del CLAD*, Ciudad de Guatemala.
- Perry, James L. et alli (2009). "Back to the Future? Performance-Related Pay, Empirical Research, and the Perils of Persistence". *Public Administration Review* 69(1):1-31, January-February.
- Trosa, Sylvie (2001). *Gestão pública por resultados. Quando o Estado se compromete*. Rio de Janeiro: Revan; Brasília: ENAP.
- Van Dooren, Wouter (2008). "Performance indicators: a wolf in sheep's clothing?"; mimeo.