



**IDÉIAS SOBRE O POLICIAMENTO NORTE-AMERICANO**  
**Número 6, Agosto de 2004**

---

**Departamentos de Polícia como Laboratórios de Aprendizagem**  
**por Edward R. Maguire**

---

*Idéias sobre o Policiamento Norte-Americano (Ideas in American Policing)* apresenta análises e comentários de importantes criminologistas sobre questões de interesse para estudiosos, profissionais e autoridades. Os trabalhos aqui publicados pertencem a uma série de conferências sob o mesmo título publicadas pela Police Foundation. As opiniões expressas neste documento são de responsabilidade do autor e não refletem necessariamente a posição oficial da Police Foundation.

© 2004 Police Foundation e Edward R. Maguire. Todos os direitos reservados.

---

**Edward R. Maguire** é professor adjunto do programa de Administração da Justiça da Universidade George Mason.

---

# **Departamentos de Polícia como Laboratórios de Aprendizagem**

**Edward R. Maguire**

Como qualquer organização, os órgãos policiais coletam, armazenam, processam, analisam e interpretam informações, tomando as providências necessárias com relação a elas (Choo 1995). Mas, muitas vezes, os órgãos não conseguem coletar ou organizar informações úteis para avaliar seu desempenho. Mesmo quando tais informações estão disponíveis, os órgãos raramente as analisam com profundidade, aprendem algo com elas e implantam mudanças com base nesse aprendizado. Na ausência de fontes sistemáticas de informação sobre desempenho, os órgãos policiais, como muitos outros órgãos públicos, geralmente não têm condições de afirmar com precisão de que forma seu desempenho evoluiu ou de que forma esse desempenho pode ser comparado com o de outros órgãos, particularmente aqueles localizados em comunidades semelhantes. Na falta de informações corretas e de capacidade para transformá-las em medidas aplicáveis, os órgãos policiais tendem a padecer do que Peter Senge (1990), guru da Administração de Empresas, chama de "incapacidade de aprendizagem organizacional". Embora em graus variados, essa incapacidade é de certo modo sistêmica, afetando todo o setor de policiamento. Neste artigo, apresentarei uma sistemática com a qual os órgãos policiais poderão superar essas dificuldades e utilizar melhor as informações e as medições para aperfeiçoar sua capacidade de aprendizagem organizacional sistemática. Quando começarem a funcionar como laboratórios de aprendizagem, as organizações policiais poderão se tornar “organizações inteligentes”, capazes de reflexão, adaptação e renovação contínuas.

## **Informação**

“Grande parte dos dados de que uma organização precisa para monitoramento e análise das informações já existe na própria organização. Infelizmente, as informações são fragmentadas e dispersas, e as pessoas que as detêm geralmente não têm consciência de seu valor para a organização nem da necessidade de compartilhá-las.” (Choo 1995, 195).

Imagine a administração de um departamento de polícia, ou de qualquer outro tipo de organização, como um sistema gigantesco de processamento de informações bem parecido com o cérebro humano. Volumes enormes de informação chegam ao cérebro a partir de inúmeras fontes. Algumas das fontes são internas, como comentários de funcionários e registros policiais, e outras são externas, como reclamações de cidadãos ou ações judiciais. Algumas são locais, como cobertura da mídia; outras, distantes, como relatórios de equipes de credenciamento profissional. Algumas são formais e estruturadas, como estatísticas de órgãos oficiais; outras, informais e não estruturadas, como as recebidas em reuniões com a comunidade. Embora todas essas fontes dêem à organização *feedback* sobre seu desempenho, as informações não chegam organizadas e bem formatadas, de forma a poder dar uma direção clara à organização. O imenso volume de informações excede a capacidade cognitiva de qualquer pessoa de processá-las com cuidado e provavelmente supera a capacidade coletiva de processamento sistemático e cuidadoso da organização. Sem um sistema avançado de gerenciamento das informações, grande parte delas provavelmente não é arquivada ou mantida em um formato que possibilite a pesquisa ou recuperação de dados. Além disso, grande parte das informações é vaga, conflitante, tendenciosa, dispersa ou difícil de interpretar. A era da eletrônica possibilitou às organizações policiais armazenar e recuperar mais informação do que nunca, mas muito pouco dela, pelo menos em forma bruta, não processada, tem utilidade para medir o desempenho da organização ou planejar

seu futuro estratégico. É um caso clássico de sobrecarga de informação.

Para criar as bases para a aprendizagem organizacional sistemática, os órgãos policiais precisam desenvolver duas novas capacidades. Primeiro, precisam melhorar a capacidade de processar, interpretar e entender as informações que já estão disponíveis, assim como de aprender com elas. Uma forma importante de fazê-lo é começar transformando dados brutos em medidas de fácil compreensão para uma variedade de dimensões do desempenho. Em outras palavras, uma estratégia de medição bem planejada e tecnicamente sólida possibilitará à organização extrair mais significado dos dados disponíveis. Até certo ponto, essa é a idéia que fundamenta a mineração de dados, uma inovação que está se tornando popular no setor privado. Segundo, os órgãos policiais devem coletar novas informações que lhes permitam descobrir padrões e tendências vitais tanto na própria organização quanto no ambiente externo. Mais adiante darei exemplos de medidas recolhidas de várias fontes que revelam informações úteis sobre como as organizações policiais evoluem e de que forma podem ser comparadas a organizações similares.

Herbert Simon, Prêmio Nobel de Economia, ressaltou há mais de quatro décadas que gerentes e administradores têm limitações cognitivas com relação à quantidade de informações que conseguem processar (Simon 1957). Diante dessas limitações, eles às vezes precisam tomar decisões com “racionalidade limitada.” Se isso for verdade, como os teóricos organizacionais acreditam, a redução da complexidade das informações disponíveis deverá melhorar a tomada de decisão. A medição é uma forma importante de resumir grandes volumes de informação, facilitando seu processamento cognitivo.

## **Medição**

“Costumo dizer que quando se mede o que se está falando, expressando-o em

números, sabe-se alguma coisa sobre o assunto; mas, quando não se pode medi-lo, quando não se consegue expressá-lo em números, o conhecimento é parco e insatisfatório; pode ser o início de algo, mas certamente não será um conhecimento científico, qualquer que seja o assunto.” (Sir William Thomson, Lord Kelvin 1889, 73-74).

Imagine um mundo sem medidas. Em vez de usar um manguito para tirar a pressão, o médico coloca os dedos em seu braço e conclui que sua pressão está normal. Sua filha chega em casa da escola sem o boletim trimestral, mas apenas com um bilhete do professor dizendo que ela se “saiu bem”. Sem velocímetro ou medidor de gasolina em seu automóvel, você não sabe em que velocidade está ou se o combustível está acabando. Quando precisa decidir sobre a compra de ações para sua carteira, seu corretor confia nos sentimentos, na intuição, e não em uma análise profunda do desempenho daquela ação no passado. Em muitas esferas da sociedade, a medição não é apenas aceita, mas espera-se que aconteça. Ela nos dá *feedback*, nos ajuda a tomar decisões e a aprender.

Nos exemplos citados, nos pareceria estranho confiar na intuição, em histórias sem comprovação ou em impressões pessoais para tomar uma decisão. Acharíamos estranho porque, para cada exemplo, já existe uma unidade de medida aceita: a pressão sanguínea é medida em milímetros de mercúrio (mmHG), medida comum da pressão dos gases; o sistema de notas com pontuação de A a F é normalmente usado para medir o desempenho dos alunos em cada curso e se transforma em uma média geral que vai de 0 a 4.0; a velocidade é medida em quilômetros por hora; os medidores de combustível calculam a fração do produto que resta no tanque; e o desempenho das ações é medido por meio de vários índices, entre eles o índice preço/lucro.

Parece que ficamos incomodados com a idéia de medir um fenômeno quando não temos

uma unidade de medida prontamente disponível. Mas todas as unidades de medida são criações do ser humano. Por exemplo, embora o uso do termômetro seja comum agora, a temperatura não é dotada de uma unidade de medida natural. Como um cientista observou, “a temperatura é um conceito difícil de quantificar” (Lindley 2004, 98). O desenvolvimento do termômetro levou vários séculos. Até a adoção das três escalas atuais — Celsius, Fahrenheit e Kelvin —, foram propostas inúmeras escalas para medir a temperatura.

Além disso, quanto mais abstrato o fenômeno, mais incomodados ficamos com a perspectiva de medi-lo. Contudo, há mais de um século os psicólogos medem fenômenos abstratos. A psicometria, ciência que mede as características psicológicas, existe pelo menos desde 1879 (Galton 1879; Ludlow 1998). Talvez os melhores exemplos de instrumentos psicométricos sejam os testes de inteligência e os vários testes de aptidão escolar que os estudantes fazem para entrar em faculdades e universidades, mas há muitos outros, entre eles os testes de personalidade e vocacionais. Por exemplo, o *Mental Measurements Yearbook*, guia de testes psicométricos, contém resenhas de mais de 2 mil testes disponíveis no mercado, só na língua inglesa (Plake, Impara, Spies e Pale 2003). Praticamente todas as características medidas pela psicometria são abstratas e intangíveis, da depressão e inteligência à satisfação no emprego e o compromisso organizacional.

Embora a psicometria tenha dado as maiores contribuições técnicas à medição das ciências sociais, não é a única área dessas ciências que trata de medição. Uma equipe de pesquisadores que trabalha no Projeto sobre Desenvolvimento Humano de Chicago criou um conjunto de métodos estatísticos chamados de “ecometria” (Raudenbush e Sampson 1999). Da mesma forma que a psicometria mede os fenômenos em nível individual, a ecometria enfoca a medição das características de aglomerados humanos como bairros e comunidades. Métodos

semelhantes foram criados para estudar as organizações, nos quais as medidas são geralmente calculadas para indivíduos e depois agregadas para formar sistemas de pontuação para as organizações (Camp, Saylor e Harer 1997; Camp, Saylor e Wright 1999). A sociometria, área já consolidada que trata da medição das relações interpessoais em grupos, é outro ramo da medição das ciências sociais (Moreno 1960). Coletivamente, vários subcampos de medição das ciências sociais vêm se desenvolvendo há mais de um século, e novas ciências continuam a surgir.

A medição tem valor. Ajuda-nos a fazer julgamentos, tomar decisões e aprender. Contudo, às vezes preferimos não medir fenômenos importantes, especialmente quando não há uma unidade de medida natural ou consagrada ou quando o fenômeno é abstrato ou intangível. Ambas as condições — a falta de uma unidade de medida consagrada e o conceito de abstrato e intangível — aplicam-se à medição do desempenho policial.

Entretanto, as ciências da medição nos ensinam que esses problemas podem ser superados. Dificilmente eles serão superados por quem não tem competência nos aspectos técnicos da medição. Há uma longa tradição de pesquisa que demonstra que a medição — especialmente em aglomerados humanos como bairros, comunidades ou organizações — sem a devida atenção aos detalhes técnicos pode produzir resultados bastante enganosos (Raudenbush e Sampson 1999; Camp, Saylor e Wright 1999). Consideremos a equação ecométrica básica utilizada para medir as características dos bairros de Chicago. Cheia de letras gregas e inúmeros símbolos, só não espanta os estatísticos mais experientes (Raudenbush e Sampson 1999). Qualquer tentativa séria de medir conceitos abstratos, inclusive praticamente todas as iniciativas para medir o desempenho de organizações policiais, deve ser feita com muita atenção aos aspectos técnicos da medição.

Embora os métodos de medição exijam sofisticação técnica, as medidas resultantes são,

ou deveriam ser, simples e de fácil interpretação. Isso geralmente significa transformar os dados em escalas padronizadas com propriedades conhecidas. Consideremos os testes padronizados, como o Teste de Aptidão Escolar [Scholastic Aptitude Test – SAT]. Os estudantes que fazem o SAT não recebem um relatório longo, com detalhes de seu desempenho em cada item, mas um resumo simples, de fácil leitura, com a pontuação em redação e matemática em uma escala conhecida de 200 a 800, juntamente com os percentis que mostram aos estudantes a comparação de suas notas com as de outros que fizeram o teste. A base do sistema de pontuação do SAT é uma bateria de métodos estatísticos complexos, mas eles são claros para aqueles que fazem o teste. De maneira semelhante, os pesquisadores do Escritório Federal de Presídios [Federal Bureau of Prisons] usam modelos estatísticos avançados para computar as medidas de desempenho dos estabelecimentos prisionais. Eles reconhecem que o método que usam “para construir tais medidas é tecnicamente complexo, mas os usuários não precisam entender os detalhes técnicos para conseguir utilizar as informações” (Camp, Saylor e Wright 1999). A medição é difícil, mas a interpretação das medidas não deveria ser. A medição tem um papel importante nos departamentos de polícia como laboratórios de aprendizagem.

Mesmo que os órgãos policiais começassem a coletar novas informações úteis para medir o desempenho e aumentassem a capacidade interna para medição especializada, de modo que as informações existentes ou novas pudessem ser transformadas em medidas de desempenho significativas, muitos deles não teriam o ingrediente principal para a aprendizagem organizacional: um grupo de comparação que os habilitasse a interpretar as pontuações. O valor do SAT e de outros testes padronizados está na possibilidade de estabelecer comparações. A maior parte das medidas na área de policiamento não permite tais comparações. Por exemplo, imaginemos que um órgão policial implante uma nova pesquisa entre os cidadãos para medir a



amabilidade e a prontidão dos policiais. Se a equipe de pesquisa fizer o levantamento uma vez e computar medidas de amabilidade e prontidão a partir dos dados levantados, como saberá que o órgão teve um bom desempenho? Sem uma fonte de comparação, ninguém vai saber. A dificuldade da equipe me lembra um problema de matemática insolúvel. “Tenho US\$ 2 para comprar algumas bugigangas. Quantas posso comprar?” São informações importantes que faltam. Sem a capacidade de comparação, as medidas que calculamos a partir de dados brutos são muito menos significativas. É aí onde entra a medição comparativa de desempenho.

### **Medição comparativa de desempenho**

“Os líderes das organizações geralmente têm dificuldade para avaliar o desempenho de suas próprias organizações. Uma das razões é a falta de informações que os ajudem a comparar seu desempenho com relação a outras organizações que prestam os mesmos serviços” (Gormley e Weimer 1999, 30).

A medição do desempenho está em todo lugar, de escolas a diretorias de empresas, de hospitais a governos locais. Atualmente é comum medir o desempenho de funcionários, subunidades, equipes de projeto e organizações como um todo. Aqui o foco é nas medidas *comparativas* de desempenho que podem ser usadas para fazer comparações entre organizações policiais de determinado lugar ou durante um determinado período de tempo. Esquematizei um sistema por meio do qual a medição sistemática do desempenho de organizações policiais pode ser usada como recurso renovável para contínuo aperfeiçoamento do setor de policiamento (Maguire 2004). Se implantado corretamente, esse sistema de medição de desempenho pode servir de trampolim para a transformação das organizações policiais em *laboratórios de aprendizagem*, que penetrem em seu ambiente interno e externo, façam medições diagnósticas, promovam mudanças e repitam regularmente o processo.

Já escrevi uma introdução detalhada à medição comparativa de desempenho na área de policiamento para a Comissão de Credenciamento de Órgãos de Segurança [Commission on Accreditation for Law Enforcement Agencies – CALEA] (Maguire 2003, 2004). Aqui apresentarei apenas a estrutura básica e partirei da idéia central de que o desempenho policial é multidimensional. Os órgãos policiais prestam vários tipos de serviços à comunidade e tal variedade dificulta a medição do desempenho. Um órgão pode estar fazendo um tremendo trabalho com suspeitos portadores de doença mental, mas pode não estar preparado suficientemente para reagir a desastres naturais ou incidentes críticos de larga escala. Outro investe muito em tecnologias de ponta, mas suas investigações produzem índices de liberação ou condenação de presos inferiores aos de órgãos semelhantes. A idéia de que os órgãos policiais conseguem ter mais sucesso em algumas áreas e menos êxito em outras não é exclusiva da polícia. Mesmo nas empresas, onde a principal medida de desempenho são os famosos “resultados”, outras dimensões do desempenho são importantes. Escândalos corporativos recentes demonstraram os perigos de focar apenas os lucros e ignorar outras dimensões do desempenho, como práticas contábeis e empregatícias honestas e confiáveis. A idéia de que o desempenho é multidimensional é central à criação de boas medidas comparativas de desempenho.

Se o desempenho é multidimensional, que dimensões devem ser medidas no setor de policiamento? A resposta a essa pergunta conceitual é o primeiro passo crucial na medição do desempenho. Embora os estudiosos possam contribuir com discussões e debates sobre as dimensões do desempenho policial, a resposta a essa pergunta deve ser o trabalho da polícia em consulta com sua comunidade. A finalidade do meu trabalho não é sugerir as dimensões para a medição do desempenho policial, mas darei dois exemplos de outros estudiosos. O primeiro, que

está no livro de Mark Moore publicado recentemente, contém sete dimensões de desempenho policial que ele acredita devam ser medidas (Moore 2002).

**Figura 1: As sete dimensões de Moore em *Recognizing Value in Policing***

1. Reduzir a vitimização criminal
2. Responsabilizar os criminosos
3. Reduzir o medo e melhorar a segurança pessoal
4. Garantir a segurança em espaços públicos
5. Utilizar os recursos financeiros de forma justa, eficiente e eficaz
6. Usar a força e a autoridade de forma justa, eficiente e eficaz
7. Servir à comunidade / obter legitimidade com aqueles que são policiados

O segundo teve seu resumo publicado em edição anterior do *Ideas in American Policing*. Trata-se do trabalho de Stephen Mastrofski (1999), no qual ele descreve seis dimensões da qualidade do serviço policial que devem ser medidas. Observa-se que todas as seis dimensões de Mastrofski podem ser classificadas dentro da sétima dimensão de Moore. Isso mostra a rapidez com que o processo de definir dimensões de desempenho policial pode se tornar complexo. Mais adiante, voltarei ao tema da complexidade das informações na medição do desempenho.

**Figura 2: As seis dimensões de Mastrofski em *Policing for People***

- \* Atenção
- \* Confiabilidade
- \* Prontidão
- \* Competência
- \* Maneiras
- \* Imparcialidade

Esses são apenas dois esquemas possíveis para definir as dimensões de desempenho policial que merecem ser medidas. Embora ambos façam sentido, minha experiência sugere que há pelo menos tantas estruturas de medição quanto o número de pessoas que as propõem.

Selecionadas as dimensões, precisamos definir como medi-las. As medidas de desempenho tradicionais na área policial geralmente derivam de dados administrativos mantidos pelo departamento de polícia. Embora esses dados possam ser muito úteis, não devem ser a única fonte utilizada em um sistema abrangente de medição do desempenho. Há várias alternativas, entre elas:

- *Pesquisas na comunidade.* Essas pesquisas se baseiam em amostras aleatórias da comunidade, incluindo pessoas que tiveram ou não contato com a polícia. As perguntas tendem a se concentrar em alguns aspectos, entre eles criminalidade e desordem percebidas, medo, histórico de vitimização e satisfação geral com a polícia. As pesquisas com a comunidade são comuns e têm valor, mas também têm algumas limitações.

- *Pesquisas sobre contato com a polícia.* Essas pesquisas são feitas com pessoas que tiveram contato recente com a polícia. As pesquisas de contato podem ser fontes vitais de informações para a administração sobre o comportamento dos policiais e as impressões dos clientes, inclusive os voluntários (por exemplo, os que chamam a polícia) e os involuntários (por exemplo, os detidos). Por exemplo, quando a polícia de Toronto fez uma pesquisa com vítimas de estupro, descobriu que elas consideraram o desempenho dos investigadores muito satisfatório, mas o trabalho dos policiais teve uma avaliação bem inferior (Rape Victims Rate Police Performance 1998 – Relatório Vítimas de Estupro Avaliam o Desempenho da Polícia 1998). As pesquisas com diferentes tipos de contatos, como detidos, vítimas de crimes ou motoristas citados pela polícia poderiam ser muito úteis — se aplicadas em várias cidades — para se saber se um órgão policial é percebido como mais ou menos justo ou eficiente do que outros. As pesquisas de contato também poderiam ser feitas periodicamente dentro de um mesmo órgão.
- *Pesquisas entre funcionários.* As pesquisas entre funcionários também servem como fontes de informação para a administração. São em geral utilizadas para explorar certos aspectos como o moral da corporação, a satisfação com o emprego, o clima social na organização e a integridade dos policiais (p. ex., Camp, Saylor e Miles 1997). Em um estudo ambicioso, os pesquisadores basearam-se nas respostas de mais de três mil policiais para criar uma medida agregada do “ambiente de integridade” em trinta órgãos policiais. Os resultados demonstraram que há muita variação entre os órgãos policiais em termos de integridade (Klockars, Ivkovich, Harver e Haberfeld 2000).
- *Observação direta.* A observação direta por pessoas treinadas é outro método

potencialmente valioso para a obtenção de medidas de desempenho. Embora a observação sistemática dos policiais de ronda seja uma atividade comum dentro dos órgãos policiais, ela não tem sido muito utilizada para comparar organizações. Contudo, há muitos exemplos de uso da observação direta para fins de comparação. Por exemplo, a Escala de Avaliação do Ambiente Escolar Infantil [Early Childhood Environment Rating Scale] utiliza observadores treinados para avaliar a qualidade das escolas de educação infantil com base na observação direta de certos fatores, tais como espaço físico e materiais utilizados, interação entre crianças e professores e várias outras dimensões do desempenho (Harms, Clifford e Cryer 1998). Em Nova York, a observação sistemática é utilizada para medir o estado de conservação de 1.078 km de rodovias em cinquenta e nove distritos (Fund for the City of New York 2002 - Fundo para a Cidade de Nova York 2001). Em Chicago, ela é usada para registrar o volume de distúrbios físicos e sociais nos bairros (Sampson e Raudenbush 1999). Embora possa produzir medidas significativas, a observação direta exige muito pessoal e, portanto, é muito cara.

*Testes independentes ou estudos de simulação.* Em vez de observar o desempenho em ambientes totalmente naturais, os testes independentes criam oportunidades artificiais para medi-lo. Por exemplo, o Instituto de Seguro para a Segurança nas Rodovias [Insurance Institute for Highway Safety] usa testes de colisão para avaliar a segurança dos veículos. Muitas empresas contratam pessoas — conhecidas como compradores secretos ou compradores misteriosos — para se fazer passar por clientes e visitar suas instalações a fim de verificar a qualidade do serviço, o trabalho dos caixas, padrões éticos e muitos outros aspectos. Nos grandes órgãos policiais, há muitos anos as unidades de

assuntos internos aplicam testes de integridade. A ABC News (2001) testou a integridade da polícia de Nova York e de Los Angeles entregando quarenta carteiras ou bolsas a policiais escolhidos ao acaso. Todas as carteiras e bolsas foram devolvidas pelos policiais com seu conteúdo intacto. Embora haja controvérsias, os testes e a simulação parecem ser um meio interessante de coletar dados sobre o desempenho da polícia de forma realmente independente.

### **Algumas decisões importantes**

Uma vez selecionadas as dimensões do desempenho, é importante resolver algumas questões sobre como utilizar os dados. Não vou descrevê-las em detalhes, mas mencionarei três delas brevemente: minimizar a complexidade das informações, considerar as diferenças de importância relativa das dimensões e garantir comparações justas.

#### *Minimizar a complexidade das informações*

Imagine por um momento que um grupo de órgãos policiais decidiu-se pelas sete dimensões de desempenho de Moore e selecionou cinco medidas de cada dimensão. Eles terão um total de trinta e cinco medidas separadas. Embora tenham reduzido parte da complexidade das informações, conforme discutido anteriormente, ainda terão um volume de dados tão grande que forçará os limites de sua capacidade de processamento cognitivo. A idéia de que a capacidade do ser humano de processar informações é limitada é um princípio fundamental da psicologia cognitiva (Miller 1956). Ao mesmo tempo, recentes avanços no estudo do processamento de informações demonstram que o modo como os dados são apresentados pode afetar nossa capacidade de processá-los. As pesquisas mostram que a apresentação de dados de forma eficiente, por meio de diagramas, gráficos e outras técnicas de visualização, pode aumentar a capacidade de processamento cognitivo (Card, MacKinley e Schneiderman 1999;



Larkin e Simon 1987; Tufte 1997; Wainer 1997). Em suma, as evidências sugerem que o número total de medidas de desempenho deve ser pequeno o suficiente para não exceder os limites de nosso processamento cognitivo. E, simultaneamente, podemos aumentar o número de medidas de desempenho processadas se atentarmos para o modo como apresentamos os dados, utilizando apresentações visuais simples e intuitivas em vez de tabelas complexas com dezenas ou centenas de números. A idéia de melhorar a cognição por meio de resumos gráficos já foi proposta em diversas áreas importantes, inclusive para prontuários médicos (Powsner e Tufte 1994), dados psiquiátricos (Powsner e Tuft 1997) e pontuação de testes de estudantes (Wainer 2002).

Uma outra opção para reduzir a complexidade das informações consiste em gerar resultados compostos que combinem os itens de cada dimensão em uma única pontuação. Quando um estudante se submete ao SAT, ao exame GRE para curso de pós-graduação (Graduate Record Examination - GRE) ou a qualquer outro teste similar padronizado, sua pontuação geral é a soma de todas as questões do teste. Esse total é padronizado para cair dentro de determinada faixa, por exemplo, 200-800 no SAT. Ninguém está muito interessado no desempenho do aluno em questões individuais, mas sim em sua pontuação geral dentro de cada dimensão, como matemática e redação. Os especialistas em medição podem usar métodos similares para criar pontuações compostas padronizadas a fim de medir o desempenho da polícia em cada dimensão. Levará tempo para a ciência amadurecer, mas o desenvolvimento de pontuações compostas padronizadas para medir o desempenho da polícia é uma possibilidade real.

#### *Considerar as diferenças de importância relativa das dimensões*

Algumas dimensões do desempenho são mais importantes do que outras. Uma reclamação comum sobre os sistemas de medição de desempenho é que eles tratam todas as

medidas da mesma forma. Isso é aceitável desde que os diferentes aspectos de desempenho sejam equivalentes, mas se alguns forem muito mais importantes do que outros, poderá haver alguma confusão. Às vezes é conveniente atribuir pesos maiores a determinadas medidas na hora de calcular a pontuação composta. Há uma variedade de métodos para calcular esses pesos, mas todos exigem conhecimento técnico. O maior desafio é atribuir os pesos de uma maneira que não seja totalmente arbitrária. Como quantificar as diferenças de importância entre múltiplas metas? Se elas forem irrelevantes, poderão ser ignoradas. Se forem importantes, tais como a amabilidade da equipe do hospital *versus* a taxa de mortalidade, então precisarão ser consideradas dentro do sistema de medição de desempenho (Gormley e Weimer 1999). Um índice de criminalidade baixo é mais importante do que tratar os cidadãos de maneira justa? Em caso afirmativo, é quantas vezes mais importante, duas? Três vezes mais importante? Um meio de determinar o peso relativo das diferentes dimensões do desempenho é realizar grupos de discussão e pesquisas entre os cidadãos. Refletir sobre a importância relativa das diferentes dimensões do desempenho da polícia também é um exercício útil que pode ajudar uma instituição policial a definir suas metas.

### *Garantir comparações justas*

Para que a medição comparativa de desempenho funcione, é preciso investir em comparações justas. Os órgãos policiais estão localizados em ambientes muito diferentes, e essas diferenças precisam ser levadas em conta durante as comparações. Essa preocupação em fazer comparações justas entre as organizações não é exclusiva do setor de policiamento. Assim como os departamentos de polícia localizados nas comunidades mais pobres e menos organizadas podem ter os maiores índices de criminalidade, os hospitais que admitem pacientes de alto risco podem apresentar as taxas de mortalidade mais altas; e as prisões que admitem os piores

bandidos podem ter as maiores taxas de reincidência. Contudo, podemos controlar essa variabilidade nos dados das organizações quando medimos seu desempenho.

Há dois métodos principais para controlar as variações nos dados: estratificação (formando grupos de pares) e cálculo de medidas de desempenho ajustadas ao risco. A estratificação é mais fácil. Consiste em formar grupos de órgãos similares. É feita uma comparação de todos os órgãos do grupo. Um meio de reduzir o constrangimento ou medo seria estabelecer um limite mínimo para o tamanho do grupo, por exemplo, dez órgãos. Todos os órgãos do grupo saberiam os nomes uns dos outros. No entanto, embora cada órgão recebesse uma classificação em comparação com os demais, nenhum teria acesso à classificação individual dos outros. Assim, por exemplo, somente o último classificado saberia que ficou em último lugar.

Embora seja o método mais fácil, a estratificação tem suas armadilhas. Alguns órgãos e algumas jurisdições são tão singulares que fica difícil classificá-los em um grupo de pares. Por exemplo, recentemente conversei sobre grupos de pares com o chefe do Serviço de Polícia de Nishnawbe-Aski, uma nação indígena canadense. Com apenas noventa e três policiais e uma jurisdição do tamanho da França, ele tem alguns destacamentos que só podem ser alcançados por via aérea. Encontrar um grupo de pares adequado a essa instituição provavelmente será uma tarefa difícil.<sup>1</sup> Além disso, às vezes os fatores que dificultam a formação de um grupo adequado podem ser sutis ou imprevisíveis. Por exemplo, algumas “*edge cities*” (Garreau 1991) podem não ter uma grande população, mas sua proximidade de grandes áreas urbanas significa que elas enfrentam problemas que as tornam peculiares em comparação com outras comunidades do mesmo tamanho.<sup>2</sup>

Calcular as medidas ajustadas ao risco é um método mais complicado, pois utiliza

controles estatísticos para ajustar as medidas às variações nos dados. Esse processo requer conhecimento técnico e um investimento substancial em testes e calibrações para assegurar que os procedimentos de ajuste ao risco sejam defensáveis do ponto de vista científico. E mais, como os índices de criminalidade ajustados ao risco são baseados na hipótese implícita de que as características demográficas e estruturais — como pobreza, raça e região — têm influência na criminalidade, os procedimentos de ajuste ao risco podem gerar controvérsias.<sup>3</sup> A base científica desse processo no setor de policiamento ainda não foi devidamente estabelecida. Em sua monografia *Ideas in American Policing*, Lawrence Sherman (1998) propôs tratar as medidas de criminalidade ajustadas ao risco como medidas de desempenho das instituições policiais. À época, ele apresentou um exemplo hipotético de taxas de homicídio ajustadas ao risco. Desde então, uma equipe de pesquisadores estabelecida na Universidade do Estado da Geórgia, Universidade Carnegie-Mellon, Universidade de Missouri-St. Louis e Consórcio Nacional de Pesquisa sobre Violência começou a desenvolver essas medidas.<sup>4</sup> De acordo com suas estimativas preliminares, apenas cinco fatores demográficos e socioeconômicos respondem por mais de dois terços da variação nas taxas de homicídio das cidades. Esse resultado inicial sugere que responsabilizar a polícia pelos índices de criminalidade sem ajuste ao risco é o mesmo que responsabilizá-la por fatores sociais e econômicos que estão muito além de seu controle.

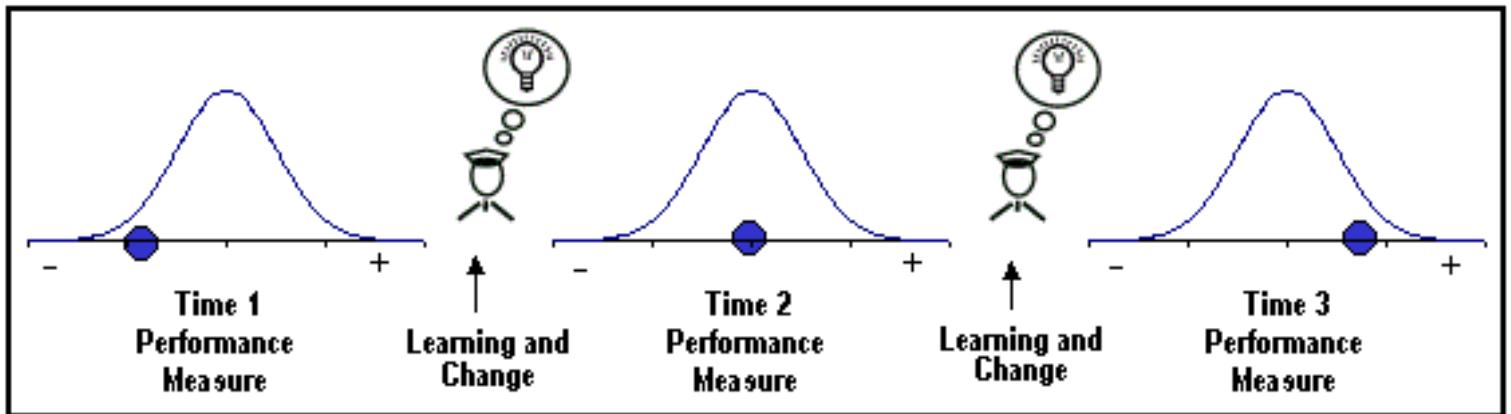
### **Por que essa abordagem pode funcionar**

Há muitas maneiras de reformar as organizações policiais; a minha visão é apenas um método que tem potencial para melhorá-las. Todas as organizações são capazes de auto-aprendizado, adaptação, ajuste, experimentação e inovação. Mas, para isso, elas precisam de informações e *feedback*. As medidas comparativas de desempenho fornecerão os dados

essenciais: como algumas organizações policiais estão se saindo em relação a outras em vários aspectos e como estão progredindo em relação a seus próprios níveis anteriores de desempenho. A pesquisa em ciências da informação indica que as organizações que se empenham em um melhor “monitoramento do ambiente” têm um desempenho melhor (Choo 1995). Até agora, essa pesquisa foi feita em empresas do setor privado e organizações sem fins lucrativos. O esquema de medição de desempenho descrito aqui representa apenas uma forma de monitoramento do ambiente de trabalho. Também há provas de que classificar as organizações pode incentivar aquelas que não estão apresentando um bom desempenho a implementar algumas mudanças (Gormley e Weimer 1999).

A teoria aqui é muito simples: a comparação motiva mudanças. As medidas comparativas de desempenho podem induzir mudanças em organizações que, se não fossem as medidas propostas, provavelmente nem saberiam que apresentam fraco desempenho. Imagine um processo pelo qual uma organização, ao descobrir que seu desempenho é mais fraco que o de outras similares, começa a implementar mudanças para melhorá-lo. Se o conseguir, deixará de ser a pior. Se esse processo de aprendizado e mudança continuar, o resultado será a melhora contínua de todo o setor de policiamento. A Figura 3 ilustra esse processo. Cada ciclo de medição de desempenho gera aprendizado e mudanças dentro da organização, resultando na melhora do desempenho.

**Figura 3: Ciclo de aprendizado organizacional**



**Time 1 = Tempo 1**

**Performance Measure = Medida de desempenho**

**Learning and Change = Aprendizado e mudança**

**Time 2 = Tempo 2**

**Time 3 = Tempo 3**

Muitas organizações policiais não sabem como anda seu desempenho em relação às demais. Mas existem alguns indicadores que podem ser suficientes para induzir mudanças nas organizações de polícia. Processos judiciais, publicidade negativa, decisões judiciais e determinações judiciais deferidas (*consent decrees*) são os motores que impulsionam as reformas, mas esses catalisadores só são ativados quando o desempenho já degenerou ao extremo. Os administradores de polícia necessitam de indicadores simples, claros e fáceis de entender, que meçam o desempenho regularmente em todo o seu espectro: negativo, neutro e positivo.

## **Por que essa abordagem pode não funcionar**

Há muitas razões para se acreditar que a medição de desempenho comparativa talvez nunca seja implementada ou talvez não melhore o aprendizado organizacional mesmo se implementada. Explorar todas as possíveis armadilhas exige um artigo inteiro dedicado só a isso. Aqui mencionarei brevemente apenas algumas delas.

- Muitas instituições podem não querer participar, principalmente aquelas que têm o pior desempenho.
- Há inúmeros exemplos de medidas de desempenho que geram comportamentos indesejados dentro das organizações, como falsificar ou manipular dados, ou permitir que aspectos do desempenho não medidos sejam afetados por causa de um foco míope naqueles que são medidos. Por exemplo, os guardas de segurança de uma usina nuclear surpreenderam todo mundo ao repelir uma série de ataques terroristas simulados. Uma investigação posterior revelou que eles haviam sido alertados quanto ao momento das simulações (Departamento de Energia dos EUA 2004).
- À medida que a ciência da medição de desempenho se desenvolve, é provável que algumas organizações reajam criticando a validade das medidas. Mais uma vez, há inúmeros exemplos disso em outros domínios da polícia, especialmente na medição do desempenho de hospitais (Iezzoni 2003). Embora existam evidências de que o monitoramento ambiental e as medidas de desempenho comparativas possam induzir mudanças em outros setores, não há provas de que elas produzam tais efeitos na polícia.
- A nova ciência da medição de desempenho pode ser facilmente contaminada ou dominada por influências políticas.
- Por fim, o que descrevi aqui supõe que exista um grupo de pessoas capazes de

transformar dados brutos em informações úteis à polícia. Essa hipótese não é verdadeira no momento, mas as universidades podem facilmente criar programas de graduação ou certificação para desenvolver pessoas com essas habilidades uma vez que a demanda por tal conhecimento começa a superar a oferta.

### **Implementação da visão**

Essa visão não pode ser implementada da noite para o dia. Na verdade, levará tempo e exigirá investimento e apoio de várias fontes importantes.

Em primeiro lugar, obviamente, exigirá a firme participação dos líderes policiais. O treinamento que dei sobre medição de desempenho a centenas de administradores de polícia e gerentes de instituições de credenciamento em conferências sobre segurança pública nos Estados Unidos e Canadá nos últimos dois anos convenceu-me de que existe uma demanda crescente por instrução e treinamento sobre essas idéias.

Segundo, o desenvolvimento da ciência de medição de desempenho exigirá auxílio financeiro e logístico por parte do governo. O Instituto Nacional de Justiça nos EUA, o Home Office (Ministério do Interior) no Reino Unido e outros órgãos governamentais similares precisarão garantir o investimento em métodos científicos para o desenvolvimento dessas medidas de desempenho. É preciso dar atenção especial ao teste, à calibração e ao aperfeiçoamento das medidas. Assim como o Serviço de Testes Educacionais [Educational Testing Service] se esforça para continuar criando e aperfeiçoando testes padronizados, esse será um processo de desenvolvimento contínuo e incremental.

Terceiro, a adoção sistemática das medidas de desempenho comparativas provavelmente criará um novo nicho ocupacional para pessoas que tenham conhecimento em pesquisa, análise



de dados, medição e visualização de informações. A visão que descrevi aqui exigirá a criação de um novo cargo civil nas organizações policiais que tenham recursos para tal. A pessoa qualificada seria um especialista em medição de desempenho que pudesse garantir que os dados fossem coletados, analisados, resumidos e relatados da forma apropriada. Talvez essa pessoa também pudesse ter conhecimento em auditoria para verificar os tipos de dados utilizados para criar as medidas. Assim como o desenvolvimento de mapas do crime levou à criação do cargo de analista criminal nas instituições policiais nacionais, o desenvolvimento de medidas de desempenho levará à contratação de pessoas especializadas em sua implementação.

Quarto, as universidades precisarão criar programas de ensino para desenvolver habilidades de medição de desempenho e fornecer certificados aos especialistas.

Por fim, como muitas organizações policiais pequenas não teriam condições de contratar seus próprios gurus para medição de desempenho, o governo teria de prestar assistência técnica a algumas delas. Talvez os institutos regionais de policiamento comunitário possam oferecer seus serviços às organizações da região que sejam muito pequenas ou não tenham condições de medir seu próprio desempenho.

## **Conclusão**

As idéias aqui apresentadas são ambiciosas. Não são baratas nem fáceis nem rápidas, mas prometem melhorar o policiamento. O primeiro passo é iniciar estudos piloto. A Comissão de Credenciamento de Órgãos de Segurança iniciou um projeto com cerca de uma dezena de organizações policiais. Seu objetivo é verificar se é possível obter medidas uniformes em várias dimensões do desempenho policial. Esse processo não será fácil, mas em termos de segurança e justiça nas comunidades, creio que vale o esforço.

## Referências bibliográficas

- ABC News. 2001. Testing police honesty: A Primetime investigation with lost wallets. Primetime, Thursday, May 17. ABC News: New York.
- Camp, Scott D., Saylor, William G., & Harer, Miles D. 1997. Aggregating individual-level evaluations of the organizational social climate: A multilevel investigation of the work environment at the Federal Bureau of Prisons. *Justice Quarterly*, 14(4), 739-761.
- Camp, Scott D., William G. Saylor, and Kevin N. Wright. 1999. Creating performance measures from survey data: A practical discussion. *Corrections Management Quarterly*, 3(1): 71-80.
- Card, Stuart K., MacKinley, Jock D. & Schneiderman, Ben. 1999. *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufmann Publishers, Inc.
- Choo, Chun Wei. 1995. *Information management for the intelligent organization: The art of scanning the environment*. Medford, NJ: American Society for Information Science and Technology.
- Fund for the City of New York. 2001. How smooth are New York City's streets? Center on Municipal Government Performance.
- Gallagher, Catherine, Maguire, Edward R., Mastrofski, Stephen D., & Reisig, Michael D. 2001. The public image of the police: Final report to the International Association of Chiefs of Police. Manassas, VA: George Mason University, Administration of Justice Program.
- Galton, Francis. 1879. Psychometric experiments. *Brain: A Journal of Neurology*. 2: 149-162.
- Gardner, Howard. 1983. *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Garreau, Joel. 1991. *Edge city: Life on the new frontier*. New York: Doubleday.
- Gormley, William T., Jr., and David L. Weimer. 1999. *Organizational report cards*. Cambridge, MA: Harvard University Press.
- Harms, T., R.M. Clifford, and D. Cryer. 1998. *Early childhood environment rating scale*. New York: Teacher College Press.
- Iezzoni, Lisa I. 2003. *Risk adjustment for measuring healthcare outcomes*. Third Edition. Chicago: Health Administration Press.
- Klockars, Carl B., Sanja Kutnjak Ivkovich, William E. Harver, and Maria R. Haberfeld. 2000. The Measurement of Police Integrity. Research in Brief. Washington, DC: National

Institute of Justice.

- Larkin, Jill, and Simon, Herbert A. 1987. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, Vol. 11, No. 1, pp. 65-99.
- Lindley, David. 2004. *Degrees Kelvin: A tale of genius, invention, and tragedy*. Washington, DC: National Academies Press.
- Ludlow, Larry H. 1998. Galton: The first psychometrician? *Popular Measurement*, 1(1): 13-14.
- Maguire, Edward R. 2003. Measuring the performance of law enforcement agencies, Part 1. *CALEA Update*, 83. <http://www.calea.org/newweb/newsletter/No83/measurement>.
- Maguire, Edward R. 2004. Measuring the performance of law enforcement agencies, Part 2. *CALEA Update*, 84. <http://www.calea.org/newweb/newsletter/No84/maguirepart2.htm>
- Mastrofski, Stephen D. 1999. *Policing for people*. Ideas in American Policing. Washington, DC: Police Foundation, (March).
- Miller, George A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81-97.
- Moore, Mark H. 2002. *Recognizing value in policing: The challenge of measuring police performance*. Washington, DC: Police Executive Research Forum.
- Moreno, Jacob Levy. 1960. *The sociometry reader*. Glencoe, IL: The Free Press.
- Plake, Barbara, James C. Impara, Robert A. Spies, and Barbara S. Pale, eds. 2003. *The fifteenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.
- Powsner, Seth M. and Tufte, Edward R. 1994. Graphical summary of patient status. *The Lancet*, 344: 386-389.
- Powsner, Seth M. and Tufte, Edward R. 1997. Summarizing clinical psychiatric data. *Psychiatric Services* 48: 1458-1461.
- Rape victims rate police performance. 1998. *Toronto Star*, July 23: C3.
- Raudenbush, Stephen W. and Sampson, Robert J. 1999. Ecometrics: Toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology*, 29, 1-41.
- Sampson, Robert J., and Raudenbush, Stephen W. 1999. Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology*, 105, 603-651.

- Senge, Peter M. 1990. *The fifth discipline: The art and practice of the learning organization*. New York: Doubleday/Currency.
- Sherman, Lawrence J. 1998. *Evidence-based policing*. Ideas in American Policing. Washington, D.C.: Police Foundation.
- Simon, Herbert A. 1957. *Models of man: Social and rationale*. New York: John Wiley and Sons.
- Thomson, Sir William (Lord Kelvin). 1889. *Electrical units of measurement. Popular lectures and addresses*, Volume 1 of 3. London: Macmillan.
- Tufte, Edward R. 1997. *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
- U.S. Department of Energy. 2004. Inspection report: Protective force performance test improprieties. Washington, DC: U.S. Department of Energy, Office of Inspector General.
- Wainer, Howard. 2002. Clear thinking made visible: Redesigning score reports for students. *Chance*, 15: 56-58.
- Wainer, Howard. 1997. *Visual revelations: Graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot*. New York: Copernicus

## Observações

- 
1. Agradeço ao chefe Wesley Luloff do Serviço de Polícia Nishnawbe-Aski de Ontário, Canadá, por essa observação.
  2. Agradeço ao chefe John Douglas do Departamento de Polícia de Overland Park, Kansas, por essa observação.
  3. Agradeço por essa observação aos membros da Subcomissão de Medição de Desempenho da Comissão de Credenciamento de Órgãos de Segurança.
  4. [Http://www.cjgsu.net/initiatives/HomRates-PR-2003-08-03.htm](http://www.cjgsu.net/initiatives/HomRates-PR-2003-08-03.htm)

## ABOUT THE POLICE FOUNDATION

The Police Foundation is a private, independent, not-for-profit organization dedicated to supporting innovation and improvement in policing through its research, technical assistance, communication, and professional services programs. Established in 1970, the foundation has conducted seminal research in police behavior, policy, and procedure, and works to transfer to local agencies the best new information about practices for dealing effectively with a range of important police operational and administrative concerns. Motivating all of the foundation's efforts is the goal of efficient, humane policing that operates within the framework of democratic principles and the highest ideals of the nation.

### DIVISION OF RESEARCH, EVALUATION, & PROFESSIONAL SERVICES

Karen L. Amendola  
*Chief Operating Officer*

M. Joe Ryan  
*Director, Crime Mapping  
& Problem Analysis  
Laboratory*

Greg Jones  
*Research Associate*

Vanessa Correa  
*Program Associate*

Vanessa Ruvalcaba  
*Research Assistant*

Robert C. Davis  
*Director of Research*

Edwin E. Hamilton  
*Director of  
Professional Services*

Raymond Johnston Sr.  
*Systems Engineer*

Eliab Tarkghen  
*Systems Engineer*

Tamika McDowell  
*Sr. Administrative Assistant*

Kevin Cazzolino  
*Graduate Research Associate*

### RESEARCH ADVISORY COMMITTEE

David Weisburd, *Chair*  
*Hebrew University and University of Maryland*

Edward R. Maguire  
*George Mason University*

David Klinger  
*University of Missouri-St. Louis*

Tracey L. Meares  
*University of Chicago Law School*

### BOARD OF DIRECTORS

*Chairman*  
William G. Milliken

*President*  
Hubert Williams

George H. Bohlinger III

David D. Cole

Wade Henderson

William H. Hudnut III

David B. Lewis

W. Walter Menninger

Laurie O. Robinson

Weldon J. Rougeau

Alfred A. Slocum

Andrew L. Sonner

Maria Vizcarrondo-DeSoto



1201 Connecticut Avenue, NW, Suite 200, Washington, DC 20036  
(202) 833-1460 ♦ Fax (202) 659-9149 ♦ e-mail: [pinfo@policefoundation.org](mailto:pinfo@policefoundation.org)  
[www.policefoundation.org](http://www.policefoundation.org)