

# Indicadores quantitativos: como obter, avaliar, criticar e aperfeiçoar

Piotr Trzesniak <sup>1</sup>

## RESUMO

Quantificar é inerente ao ser humano, e é inegável que trazer essa prática para a construção do conhecimento foi decisivo para o grande sucesso das ciências exatas e das engenharias já no século XIX. Assim sendo, a transformação de informações em números para entender e controlar processos foi transferida para as ciências humanas e sociais, mas... o sucesso não foi o que se esperava, levando muitos pesquisadores dessas áreas a, até, rejeitar a quantificação em geral. Embora tal postura seja compreensível, argumentamos que ela não decorre da quantificação em si, mas de que, a ela agregadas, vieram subliminarmente duas assunções bastante questionáveis e insuficientemente discutidas: as de que, como ocorre na física, nas áreas das ciências humanas e sociais, (i) os fenômenos são predominantemente determinísticos, e (ii) existe a expectativa de um “valor verdadeiro” claro e bem definido para a informação quantificada. O fato é que a quantificação nas áreas não exatas torna-se muito mais útil e eficaz caso se reconheça que (i) os fenômenos são mais bem entendidos se abordados como estocásticos, (ii) não há a expectativa de um “valor verdadeiro” para as informações quantificadas, elas são predominantemente *distribuições*, intervalos, e (iii) as formas de quantificar, embora sem dúvida totalmente válidas, são menos objetivas. Neste artigo, caracterizamos cada um desses contextos, visando a discriminá-los clara e definitivamente. Diferenciamos *medidas de indicadores*, e nos concentramos nestes últimos, classificando-os conforme suas várias características, aplicações e funções, e concluindo com um elenco de requisitos de qualidade que eles devem atender para maximizar eficácia, utilidade e precisão.

**Palavras-chave:** Indicadores. Medidas. Processos estocásticos. Quantificação. Modelagem.

## 1 INDICADORES: ALIADOS IMPRESCINDÍVEIS NA CONSTRUÇÃO DO CONHECIMENTO EM TODAS AS ÁREAS

A rejeição de qualquer tipo de quantificação por pesquisadores de diversas áreas do conhecimento é algo bastante comum. Há, sem dúvida, múltiplas razões de peso para tal atitude, mas várias delas carecem de fundamento e não resistem a uma análise aprofundada, enquanto que outras resultam de algum mal entendido, que acabou se perpetuando mesmo sem ter qualquer sustentação lógica ou factual. A parte inicial deste texto é dedicada a quebrar o falso paradigma de que a quantificação é inútil ou mesmo ruim, e que deve ser evitada. Pelo contrário: a realidade das áreas científicas mais antigas e consolidadas demonstra inequivocamente que se trata de um recurso rico e poderoso para qualquer ramo do conhecimento – embora não necessariamente para *todos* os fenômenos e processos dessa área – e *desde que* – é indispensável – se leve em conta a natureza dos fenômenos e processos respectivamente envolvidos. É um erro fatal, e uma das grandes causas da rejeição à quantificação, a simples e ingênua transposição da maneira de quantificar e de lidar com os valores de uma área para outra – especialmente das chamadas “duras” para as humanas e sociais.

<sup>1</sup> Doutor em Física. Departamento de Ciência da Informação, Universidade Federal de Pernambuco - piotrze@gmail.com

Este texto discute indicadores: como concebê-los, como utilizá-los, como criticá-los e aperfeiçoá-los, e o que deles esperar e o que não. Visa também a levar quem os rejeita a refletir sobre eles, de modo a, eventualmente, reconhecer o seu valor, seu mérito e, de fato, sua necessidade para o avanço do conhecimento. Tudo isso é uma questão de empregá-los adequadamente e de ter consciência de suas limitações. Mas aceitá-los como úteis passa também pela compreensão da natureza dos fenômenos, dos processos que ocorrem no Universo e que procuramos entender com a nossa Ciência. Referimo-nos a Universo em seu sentido pleno: tanto o Natural, estudado, entre outras, pela Física, pela Química, pela Biologia, como o Humano, objeto da Psicologia, da Sociologia, da Antropologia, também entre outras.

O Universo é dinâmico, está permanentemente em mutação, e essa se dá de modo muito complexo e intrincado. Mas a mente humana é capaz de isolar algumas partes dessa mutação, particularmente aquelas que se repetem frequentemente e cujo desfecho é de importância para a sobrevivência ou o bem-estar – esses são os *fenômenos* ou *processos*. Entendê-los, com o propósito de conduzir seu desfecho sempre no sentido de um resultado almejado *a priori* (ou seja, de gerar uma *tecnologia*) é a origem do conhecimento científico, foi sua primeira e maior preocupação, e segue sendo uma das mais importantes na atualidade.

## 2 O DETERMINÍSTICO E O ESTOCÁSTICO

Possuir conhecimento significa, então, ser capaz de modificar o comportamento do Universo, levando-o a caminhar no sentido do bem-estar da humanidade<sup>2</sup>. Por exemplo, inexistindo vacinas, ou não havendo antibióticos, o andamento espontâneo do Universo conduziria muitas vezes a uma quantidade importante de mortes prematuras. Porém a humanidade descobriu procedimentos (isto é, adquiriu *conhecimento*) que alteram essa evolução, evitando as grandes epidemias que, por mais de uma vez, devastaram populações inteiras em épocas passadas. Com algum empenho, seria possível construir uma lista de exemplos com certamente centenas de milhares, senão milhões, de soluções que ilustram essa concepção do que seja conhecimento, distribuídas pelas diversas áreas e especialidades profissionais, das mais exatas e *determinísticas*, como a física e as engenharias, às mais complexas<sup>3</sup> e de vocação *estocástica*, como a psicologia científica e clínica – passando por todas as muitas outras a que a mente humana decidiu se dedicar.

A fluência do texto e das ideias discutidas até o final do parágrafo anterior conduziram naturalmente a uma diferenciação entre as áreas do conhecimento, conforme nelas predominem – sem excluir a possibilidade de que convivam – processos determinísticos e estocásticos. Compreender essa diferença é sem dúvida a questão mais importante para navegar com solidez e bom discernimento no espaço das medidas e dos indicadores. Uma das principais causas de rejeição à quantificação advém de se tentar transpor ingenuamente a processos estocásticos os métodos, técnicas e procedimentos adequados aos determinísticos. Isso é absolutamente impraticável e, muito mais do que contribui para, *retarda* o avanço do conhecimento.

Torna-se então *absolutamente fundamental* para bem compreender e usar indicadores e medidas reconhecer a existência e as diferenças entre três tipos de fenômenos, com base na natureza de sua relação causal.

Ao primeiro e provavelmente majoritário tipo pertencem os fenômenos cuja observação não permite inferir qualquer regularidade, repetição ou influência de condições causais. Esses ficam, então, fora do alcance da Ciência. Os outros dois tipos – o determinístico e o estocástico – são os que apresentam uma relação causal que faz sentido à lógica do cérebro humano, portanto nos permitem construir conhecimento.

Um físico faz sua experiência e obtém o desfecho. Repete-a uma, duas, cem, milhares de vezes, e o desfecho é sempre o mesmo. Varia a intensidade de certas características necessárias da situação causal e observa – recorrentemente – variações correspondentes na intensidade de características do desfecho, mas *não* na sua natureza. Então tudo aponta para tratar-se de um processo determinístico.

Um empresário, por sua vez, cria um benefício para seus colaboradores de um determinado setor e todos ficam muito felizes e mais produtivos. Entusiasmado, estende o benefício para todos os setores da empresa e... metade dos colaboradores fica descontente e revoltada, 30%, indiferentes e apenas 20% ficam mais contentes e mais produtivos. Bom, ele se deparou com um fenômeno estocástico. Desfechos de naturezas distintas, até conflitantes, apesar de idênticas situações causais.

O Quadro 1 destaca várias diferenças entre os processos determinísticos e estocásticos<sup>4</sup>, elaborando a distinção entre elas. Mas o aspecto que nos parece mais fundamental e de imediata clareza para quem

<sup>2</sup> Infeliz e paradoxalmente, mentes perturbadas também promovem a construção de conhecimento visando dizimar populações, desenvolver procedimentos, máquinas e artefatos que matam, mutilam e inutilizam (veja esse tema informalmente discutido em <http://www.youtube.com/watch?v=2jiWgsulp8w>). É triste reconhecer que tal possa fazer parte da natureza humana, mas o justo desejo e a expectativa é que se trate de práticas definitivamente extintas ou em acelerada extinção, e que o entendimento de todos em favor do bem-estar prevaleça na nova era.

<sup>3</sup> Não, não se trata de um equívoco: as ciências ditas "exatas", especialmente no que concerne aos fenômenos determinísticos, harmonizam-se naturalmente à álgebra e ao cálculo diferencial e integral convencionais, o que faz muito simples a compreensão e a modelagem dos respectivos processos e possibilita a preparação de projetos eficazes e bem-sucedidos. O grande desafio para o conhecimento humano é compreender e modelar os fenômenos das áreas de humanas, e parece já estar bastante claro que a solução para isso não é a modelagem matemática clássica das ciências exatas.

pesquisa é o seguinte: considere um grande número de repetições de um mesmo fenômeno, que se desencadeia a partir de uma determinada situação causal conhecida.

- Nos *determinísticos*, o desfecho é sempre o mesmo.
- Nos *estocásticos*, há um conjunto limitado de desfechos diferentes, cujo aparecimento, no entanto, obedece a uma distribuição de frequências fixa. Assim, se os desfechos observados forem A, B e C, poder-se-á ter (por exemplo) A em 35% das repetições, B em 52% e C em 13%.

Quadro 1 - Caracterização das relações causais determinísticas e estocásticas

	<b>Processo determinístico</b>	<b>Processo estocástico</b>
Quantidade de condições necessárias na situação causal	Poucas, máximo em torno de dez, identificadas.	Muitas, dezenas, até centenas delas, não todas identificadas
Nas repetições do fenômeno, sem que haja alterações na situação causal	O desfecho não muda, é sempre o mesmo, bem determinado.	Há um número limitado de desfechos possíveis, que aparecem conforme uma distribuição de frequências conhecida; Apenas um desses desfechos será observado a cada repetição.
Alterando a intensidade de alguma condição necessária, qual a consequência mais comum?	Altera-se a intensidade de alguma característica do desfecho, mas não a sua natureza. A alteração é adequadamente descrita por uma expressão matemática.	O elenco de desfechos permanece, mas a frequência associada a cada um pode mudar. Também pode mudar a intensidade de alguma característica do desfecho que vier a se concretizar.
Adicionando-se uma condição relevante qualquer ao conjunto das necessárias	A natureza do desfecho muda, tem-se outro processo que, em princípio, será determinístico.	Podem-se mudar apenas as frequências associadas aos vários desfechos, mas também ter modificações no conjunto de desfechos possíveis.
Validade	O processo determinístico é refutado se, mantidas todas as condições, o desfecho não aparecer uma única vez.	A ocorrência de um dado desfecho num caso não significa que não possa ocorrer o contrário em outro, desde que esse "contrário" faça parte dos desfechos possíveis; nada se pode concluir com base em apenas um caso isolado.

Fonte: elaborado pelo autor (2014)

Em termos práticos, isso significa que, se um processo é estocástico, o agente pode fazer tudo exatamente do modo como fez na vez anterior e obter um resultado completamente diferente, até mesmo oposto. Estar ciente de que se está pesquisando ou intervindo em um fenômeno estocástico é absolutamente fundamental para pensar melhor, considerar alternativas, ficar mais atento e preparar-se para agir eficazmente em caso de desfechos distintos do almejado.

Vamos complementar nossa discussão acerca do determinístico e do estocástico com quatro comentários, o primeiro dos quais muito prático e extremamente relevante, para o qual recomendamos, então, um estudo mais atento. Os outros três envolvem aspectos mais sutis, importantes e interessantes, mas não tão vitais.

1. É bastante comum, em certos meios, haver a tentativa de manipular a opinião de ouvintes e interlocutores, e induzi-los a ficarem confusos, argumentando *deterministicamente* com relação a processos *estocásticos*. Adversários de cintos de segurança citam exemplos de situações em que os envolvidos em colisões morrem mesmo com o equipamento colocado, fumantes se justificam citando o *avô com mais de noventa anos que desde os quinze consome dois maços por dia, enquanto o primo (que nunca pôs na boca qualquer cigarro ou similar) faleceu com menos de vinte*. Estas situações somente derrubariam as relações entre uso do cinto e redução de mortes no trânsito, e entre fumo e doenças, caso tais relações fossem determinísticas. Elas são, no entanto, estocásticas, e, assim, os exemplos descritos estão absolutamente dentro da normalidade. Tratando-se de probabilidade, pouco peso ou sentido existe no evento isolado, a partir dele nada se pode provar ou invalidar. Conclusões exigem estudos e comparações em *conjuntos representativos* de situações de igual natureza.

2. O Universo – natural ou humano – não é estocástico ou determinístico. Essas são apenas duas maneiras de tentar compreender o seu comportamento de um modo que faça sentido para a nossa mente e não têm a mínima pretensão de entrar no mérito de como o Universo é de fato.

3. Existe um argumento de que tudo é na verdade determinístico, o problema é que não se conhecem

<sup>4</sup> A descrição de processos estocásticos aqui apresentada está longe de ser rigorosa, e nem o pretende. Destina-se apenas a chamar a atenção para seu eventual emprego para a modelagem nas áreas de ciências humanas e sociais aplicadas. Um aprofundamento iria além do que este trabalho se propõe a fazer.

todas as variáveis envolvidas e, portanto, não se é capaz de escrever a expressão matemática descritiva do fenômeno. Essa forma de ver coloca o determinismo como o modelo absoluto e hegemônico, mas não traz qualquer vantagem prática: se as variáveis e expressões são desconhecidas, então “não se conhece o conhecimento”... Se for para confrontar, faz muito mais sentido dizer que o modelo *estocástico* é hegemônico – não apenas porque a quantidade de fenômenos a que ele se aplica é muito maior, mas também porque o determinístico pode ser visto como um estocástico de um único desfecho, que ocorre com frequência 100%.

4. Dependendo de quanto o elenco de desfechos é alterado em um processo estocástico, há ao menos duas possibilidades de modelagem: dizer que se trata de um novo processo ou considerar que o processo é o mesmo, mas que alguns desfechos não eram observados antes porque sua frequência de ocorrência era zero (ou que deixaram de ser observados na nova configuração causal porque sua frequência tornou-se zero). A visão que irá prevalecer será sempre a mais conveniente, mais simples e mais lógica.

### 3 QUANTIFICAR INFORMAÇÕES

A quantificação, então, é um recurso de que se deve lançar mão sempre que possível – mas, embora importante e útil, *ela somente é pertinente quando associada a um conceito qualitativamente claro e bem estabelecido*. Pode-se, e faz sentido, ter um conceito qualitativo que não se pode quantificar. O inverso, porém, não é válido: uma quantificação somente passa a fazer sentido quando fica claro e bem definido a que conceito (qualitativo) ela se refere. Observe-se que é possível, por algum tempo, existir alguma insegurança relativamente a o valor obtido efetivamente dizer respeito ao conceito em estudo. Nesse período, o que está sendo pesquisado é a própria quantificação, ela é insegura, e o uso de seu resultado em outras pesquisas e intervenções deve ser feito com reserva e cautela. Apenas quando a correspondência valor-conceito estiver solidamente estabelecida é que seu emprego poderá ocorrer sem maiores preocupações.

Nessa fase, em que a quantificação de um conceito está sendo pesquisada, costumam surgir muitos trabalhos aparentemente científicos, mas que não apenas em nada contribuem para o avanço do conhecimento, como, de fato, o prejudicam. Trata-se de uma pessoa que aproveita um construto controverso, mas que esteja em evidência, monta um questionário exclusivamente de sua própria cabeça, aplica-o a uma população, faz uma estatística pouco transparente (de preferência utilizando um pacote computacional famoso) e diz que *mediu o conceito naquela população*. Isso é absolutamente inútil porque: (i) um questionário apressadamente montado, por uma só pessoa e insuficientemente discutido, pode estar “medindo” algo totalmente distinto do que autores e autoras presumem, e (ii) fornecer apenas o nome de um construto ainda controverso, sem para ele explicitar um enunciado claro e rigoroso, torna impossível compor ou comparar o trabalho em questão ao de outros pesquisadores: embora todos se refiram a um mesmo nome, o significado associado a esse nome pode variar bastante de um para outro.

#### 3.1 A quantificação que leva às medidas: contagem e comparação a padrões

A primeira coisa a entender é que *Quantificar é inerente ao ser humano*. Isso se revela pela existência dos números naturais, os inteiros positivos. *Contar elementos é a forma básica de quantificar* e quase certamente se iniciou para caracterizar o tamanho de rebanhos, de cidades (contando casas e pessoas) e o total disponível de objetos, de entidades *discretas* em geral, ou seja, que apenas existem integralmente. Não é difícil concordar, por exemplo, que não existe meia ovelha: se for meia, não mais será ovelha...

Com igual naturalidade, surgiram as operações fundamentais: dois jovens que se casassem, um possuindo oito vacas e o outro 11, reuniriam um patrimônio comum de  $8 + 11 = 19$  vacas. Se três morressem, ficariam com  $19 - 3 = 16$ . Se vendessem duas vezes três vacas e, por igual valor unitário, três vezes duas vacas, apurariam um montante idêntico, ou seja,  $2 \times 3 = 3 \times 2$ . Daí saiu a aritmética e, posteriormente, a álgebra e o cálculo infinitesimal. Observe-se que o uso da aritmética permitia prever com perfeição o resultado de uma transação (como a das vacas) *sem efetivá-la*. Antecipar o desfecho de um processo e, sendo o caso, modificá-lo para que beneficie a humanidade é, de fato, a meta maior do trabalho científico.

A contagem de itens ou ocorrências discretas é uma forma de quantificação forte e objetiva, sendo largamente empregada em praticamente todas as áreas do conhecimento: contam-se itens comercializados, artigos publicados, crimes praticados, nascimentos e falecimentos, acidentes,... a lista é interminável.

Essa primeira maneira de quantificar, e o sucesso, a ela atrelado, de fazer previsões acertadas e bem definidas sobre o futuro, certamente não passaram despercebidas aos nossos ancestrais. As noções de múltiplos e submúltiplos (como *duas vezes mais* ou *cinco vezes mais*, e *metade* ou *dois terços*, respectivamente) mostravam-se úteis e importantes para a vida cotidiana. Em paralelo, porém, outros aspectos do dia a dia, igualmente muito importantes, não se prestavam a essa abordagem. Noções como, por exemplo, tamanho, massa e tempo. O problema é que essas noções não podiam ser contadas adequadamente – elas se apresentavam *contínuas*, não eram *discretas*, não tinham um elemento básico – como os rebanhos – cujas

repetições fosse possível contar. Para a terceira delas – o tempo – até se poderia pensar em usar o dia, mas, diferentemente do caso das ovelhas, fazia sentido e era necessário falar em “meio dia”, “um quarto de dia”. Um dia inteiro revelava-se um elemento básico grande demais. Além disso, parecia ser bem conveniente poder avaliar e trabalhar com essas noções usando os conceitos de dobro, metade, três quartos e, eventualmente, fazer contas com elas. Então o que cria o engenho humano? Introduz para essas noções um *elemento básico*, que denomina *unidade de medição*, e, nos casos em que tal se mostra interessante e útil, passa a contar quantas repetições da unidade “cobrem” a intensidade com que a noção está ali presente. O que é uma régua comum senão um dispositivo que conta quantos milímetros “cabem” na distância que se deseja determinar? A contagem vem até marcada ali, de 10 em 10, para ajudar! O que é uma balança senão um instrumento que conta quantos gramas ou quilogramas correspondem a um objeto qualquer? E o relógio? Não é um contador de segundos, minutos, horas?

Essa segunda forma de quantificar traz a força e a objetividade da contagem de casos discretos a situações onde o objeto de interesse é *contínuo*, isto é, em que se considera (mesmo que não seja exatamente verdade!) que é possível dividi-lo em partes cada vez menores sem que ele perca sua natureza ou identidade.

O grande sucesso da física chamada *clássica* (que se consolidou durante o século XIX) deve-se a essa quantificação e ao modo espetacular como a matemática, mesmo por caminhos totalmente divorciados da “realidade”, é capaz de antecipar, de simular, de descrever um número significativo de processos e fenômenos que ocorrem no universo. Especialmente, daqueles cuja relação causa-efeito é de natureza *determinística*, o contexto ótimo para ser descrito através de uma lei matemática típica, em que “ $2 + 2$  são sempre 4”!

Palavras de William Thomson, que posteriormente recebeu o título de Lord Kelvin, dão uma ideia do prestígio da quantificação e do determinismo por volta de 1860:

Afirmando muitas vezes que, se você medir aquilo de que está falando e o expressar em números, você conhece alguma coisa sobre o assunto; mas, quando você não o pode exprimir em números, seu conhecimento é pobre e insatisfatório; pode ser o início do conhecimento, mas dificilmente seu espírito terá progredido até o estágio da Ciência, qualquer que seja o assunto. (RESNICK; HALLIDAY, 1973, p. 1).

Evidentemente, não se pode concordar totalmente com o entusiasmo de Kelvin. Atualmente, não parece que, se você *não puder* expressar algo em números, *seu conhecimento seja pobre e insatisfatório e que o espírito não terá chegado até o estágio de Ciência*. Mas também não se deve renegá-lo sumariamente: é inquestionável que, se você *puder* expressar aquilo de que está falando em números, quase todas as vezes o seu conhecimento será mais rico e mais satisfatório. Ressalvamos desde já, porém, que essa conclusão é tudo o que se pode transpor do contexto do determinismo para as áreas das ciências humanas e sociais aplicadas. Nessas áreas, é preciso levar em conta algumas diferenças de grande envergadura, capazes de acarretar consequências indesejadas e imprevisíveis:

- As formas de quantificar são mais fracas e menos objetivas, o que implica que precisam ser concebidas e executadas com muito cuidado, de modo que o resultado numérico obtido caracterize efetiva e adequadamente o que se deseja observar. Pela mesma razão, é preciso ter cautela no emprego, uma vez que a informação pode ser insuficiente ou estar contaminada por efeitos a ela estranhos.

- A maior parte dos fenômenos e processos passíveis de uma modelagem causa-efeito são de natureza estocástica. Não atentar para essa característica, e analisá-los sob uma óptica determinística, é um absurdo equívoco, não é capaz de oferecer qualquer avanço no conhecimento. Acreditamos que essa seja uma razão importante para a rejeição ao uso de indicadores e relações matemáticas por tantos pesquisadores dessas áreas. O problema, porém, não é da quantificação, mas da tentativa de aplicá-la inadequadamente.

Um último aspecto fundamental, frequentemente negligenciado, diz respeito a realizar-se a quantificação dentro da expectativa de que exista, para o resultado, um *valor verdadeiro bem definido*, em busca do qual se está. Esse é, para ilustrar, o caso da Física e das Engenharias: a quantificação da massa de um objeto, do comprimento de uma peça, da corrente elétrica em um fio ou cabo é sempre encarada como uma “caçada” ao valor verdadeiro; repetições do procedimento conduzem a valores muito próximos entre si, sendo as diferenças atribuídas a fatores variados (ambientais, do operador, dos instrumentos utilizados), mas *não* a variações daquilo que está sendo medido. Faz-se então uma média aritmética, que é considerada a melhor expressão do valor verdadeiro e empregada como tal em cálculos e equações, gerando resultados aos quais igualmente se atribui o *status* de “valores verdadeiros”. Muitas vezes, a média é acompanhada da chamada *incerteza de medição*, um valor que somado a ela e dela subtraído define os extremos de um intervalo acerca do qual se afirma que *contém o valor verdadeiro com máxima segurança*. Existe um procedimento matemático bastante confiável para encontrar esse intervalo para um valor calculado, desde que a incerteza não supere poucas unidades por cento dos valores (obtidos por média) das variáveis dependentes. Mas o aspecto-chave é a crença de que existe para tudo um valor verdadeiro. Bem diferente é a situação que se tem quando se quantifica no contexto das ciências humanas e sociais aplicadas. Por exemplo, quando

se deseja caracterizar uma população e se observam idade, renda, número de pessoas que constituem a família e similares. Aí não se está em busca de um valor verdadeiro. Pelo contrário, reconhece-se a *priori* que o resultado é intrinsecamente uma *distribuição*. Isso muda radicalmente o significado das médias que, de legítimas representantes de valores verdadeiros, passam a ser apenas centralizadoras de distribuições, drasticamente enfraquecendo sua força e confiabilidade como variáveis independentes em cálculos. No caso em que se admite a existência de um valor verdadeiro, a média o representa bem; porém, quando tal expectativa inexistente, quando o que se quantifica é uma distribuição, a média não é uma boa representante e tomá-la como tal sem reservas é um erro. Na verdade, dever-se-ia, por exemplo, efetuar qualquer cálculo com o valor médio e com vários outros valores, a ele superiores e a ele inferiores, de modo a cobrir adequadamente toda a distribuição; o resultado, obviamente, corretamente e coerentemente, será, então, também uma distribuição.

O equívoco que pode ocorrer com relação a essas duas situações – o *segundo grande equívoco* – é raciocinar, trabalhar e discutir no espaço das distribuições como se se estivesse no de valores verdadeiros. Fazê-lo, tal como pensar deterministicamente em contextos estocásticos – o *primeiro grande equívoco* – leva a conclusões desapontadoras e a uma injustificada e insustentável descrença nas virtudes da quantificação.

### 3.2 A quantificação que leva aos indicadores: escalas, instrumentos e testes

Os procedimentos de quantificação descritos na seção anterior são os mais rigorosos, os mais metrológicos, com pouco espaço para a subjetividade, e seu resultado é costumeiramente denominado *medida*. Existem, porém, muitas circunstâncias em que aplicá-los não é possível. Empregam-se então outros métodos, em que a subjetividade é maior, acarretando mais dispersão, menor reprodutibilidade e menor exatidão nos valores numéricos resultantes. Em consequência, o grau de informação contido nesses valores é menor do que no caso metrológico e, de fato, pode variar entre relativamente pobre e bastante robusto.

A robustez reduzida, aliada a que essas quantificações não metrológicas acabam sendo aplicadas majoritariamente em fenômenos estocásticos, faz com que os resultados sejam identificados como *indicadores*, e não medidas – eles muitas vezes dão uma *indicação*, mas não uma certeza absoluta do que se pretende saber. Ressalve-se, porém, que a robustez tende a crescer com o aperfeiçoamento do processo de observação, levando a um resultado cada vez mais confiável, ao ponto de a denominação *medida* tornar-se perfeitamente adequada.

#### 3.2.1 O item-escala simples

A forma mais simples de quantificação não metrológica é a item-escala. Concretiza-se mediante a proposta de um enunciado acerca do qual é solicitado um posicionamento do respondente relativamente a uma escala numérica. Observe-se a grande causa para a menor robustez dessa quantificação, relativamente à das medidas físicas: lá, a unidade é externa e comum a todos os observadores (como o kg, o metro, o segundo); aqui, cada respondente decide, por critérios subjetivos próprios, o que é 1, o que é 2 etc., o que reduz significativamente a acurácia do resultado.

À exceção das escalas dicotômicas, com apenas duas possibilidades de resposta, é sempre recomendado que se tenha um número ímpar de valores, propiciando a opção por um ponto equidistante dos extremos. A quantidade preferencial de alternativas é cinco, oferecendo duas possibilidades de escolha acima e duas abaixo do valor central. Excepcionalmente, quando houver razões específicas significativas, pode-se trabalhar com um total de três ou de sete pontos. Não é recomendável usar mais do que sete pontos. Por exemplo, se forem nove, os respondentes tendem a não discriminar entre si tanto os níveis intermediários baixos (segundo e terceiro) como os altos (sétimo e oitavo), que acabam se confundindo respectivamente, e a escala funcionará de fato como uma de sete opções.

Muitas vezes, a escala é disfarçada, como se fosse qualitativa<sup>5</sup>, empregando-se uma valorização verbal (como *péssimo, ruim, regular, bom, ótimo*), com ou sem a apresentação explícita de números. *Quando o objetivo é quantificar*, é preciso que tal verbalização apresente *uniformidade* no espaçamento de cada nível para o seguinte, assim como um *equilíbrio* entre os níveis acima e abaixo do médio. Uma anedota a esse respeito diz que um político em campanha, concorrendo com o ocupante do cargo em disputa, encomendou uma pesquisa acerca do desempenho do rival em que os itens eram *bom, regular, ruim e péssimo*, ou seja, o segundo melhor item já era *regular...* obviamente, o resultado só poderia terminar enviesado.

A propósito da numeração da escala, recomenda-se iniciar de 1 caso ela seja exibida ao respondente, pois algumas pessoas se sentem desconfortáveis para atribuir o zero e evitam fazê-lo. A análise dos resultados,

<sup>5</sup> Uma escala genuinamente qualitativa apresenta a grande vantagem de não precisar seguir os princípios de equiespaçamento de níveis e equilíbrio de extremos, permitindo avaliações originais e criativas, como, por exemplo, discriminação somente a partir de um piso mínimo. Imagine-se uma situação em que deva ser atendido um certo número de requisitos obrigatórios para suficiência e, depois disso, outros, acessórios, para discriminação. Então pode-se construir uma escala em que, indiferentemente, a falta de atendimento de um, ou de dois ou de todos os requisitos obrigatórios corresponda a um mesmo valor, insuficiente, e, somente se todos esses obrigatórios forem atendidos, atribuir suficiente, bom e excelente com base nos acessórios. Somente a não obrigação de quantificar permite esta versátil alternativa.

porém, deve sempre ser feita iniciando a escala com zero, ou seja, se a apuração foi feita com base em 1, 2, 3, 4 e 5, deve se subtrair 1 de todos os valores obtidos. A razão é simples: o 3 está no centro da escala de 1 a 5, mas corresponde a 60%, e não 50% de 5. Isso favorece um erro de desatenção, como concluir que “a aprovação atingiu 60% do máximo” quando, de fato, ela tiver sido de 50%. Note-se que isso se ajusta se a escala for 0, 1, 2, 3 e 4.

Finalmente, as quantificações tipo item-escala têm aplicações em coletas de opinião sobre um aspecto específico, quando se consultam muitas pessoas não especialistas, ou em avaliações, em que um ou mais (poucos) especialistas julgam um contexto específico. A validade do resultado é muito diferente nos dois casos, o que se ilustra com outra anedota. A questão era se a pessoa ocupante da presidência da república ressonava ao dormir. Uma maneira de descobrir foi preparar uma amostragem popular em 15 cidades de todos os portes pelo país afora, propor a questão às pessoas em geral, tabular, analisar estatisticamente e concluir que sim (ou que não) com margem de segurança inferior a 2%. Outra é perguntar apenas ao cônjuge e aos filhos.

### 3.2.2 Instrumentos do tipo item-escala

A quantificação item-escala simples é adequada para aspectos pontuais, específicos. Quando o interesse de pesquisa se estende a contextos ou construtos mais abrangentes, a solução passa a ser usar um instrumento, constituído por um conjunto de questões tipo itens-escala. A propósito, qualquer instrumento, seja item-escala, seja prova-teste (discutido na seção a seguir) tem de ser exaustivamente testado, avaliado, aperfeiçoado e adaptado até que atinja o *status* de validação e possa ser usado com segurança e confiança.

Especialmente no caso de instrumentos item-escala extensos, costumam-se incluir *itens de atenção* (como “nesta questão, assinale a opção *regular*”) e *itens de coerência* (propor uma questão mais de uma vez, parafraseada, em pontos distintos da sequência de itens). Isso permite ter uma idéia do engajamento do respondente e, se for o caso, atribuir um peso diferenciado à sua participação (por exemplo criando três grupos:  *muito, regularmente e pouco* engajados).

Um emprego especialmente interessante de instrumentos item-escala é a análise fatorial, uma técnica estatística que permite agrupar itens em *fatores*, não correlacionados entre si, mediante a determinação da contribuição de cada um desses fatores para a variância dos dados originais. Por exemplo, num instrumento de 60 itens (eventualmente correlacionados), dedicado a estudar um dado construto, descobre-se que a melhor explicação para a variância observada é obtida agrupando-se os itens em três categorias, o que sugere ser o construto em questão uma variável dependente de três independentes.

Como no caso item-escala simples, instrumentos item-escala podem destinar-se a colher opiniões ou a realizar a quantificação de conceitos de interesse científico. Os cuidados e o rigor em todas as etapas – da criação e validação à aplicação – precisam ser especialmente observados na segunda finalidade.

### 3.2.3 Instrumentos do tipo Prova/Teste

A quinta maneira de quantificar é através de provas e testes. Embora esses instrumentos sejam também comumente constituídos por itens, diferem dos do tipo item-escala por solicitar respostas (não necessariamente verbais - ver último parágrafo da próxima seção) que ou são abertas, ou, se apresentadas aos respondentes, não se organizam segundo uma escala. Nos instrumentos do tipo item-escala, a própria resposta quantifica direta e imediatamente o aspecto a que o item se refere; nos do tipo prova/teste, é preciso considerar mais de um item para apurar o valor do conceito ou construto envolvido.

Um modo alternativo, de fato complementar, de olhar essa mesma diferenciação é considerar um construto ou conceito complexo, que não permita uma quantificação direta, implicando a necessidade de se o quebrar em múltiplos aspectos subsidiários, que se admite que o componham. Se tais aspectos forem cada um quantificável isoladamente, teremos a situação de um instrumento item-escala. Já se a quantificação somente for possível considerando-se vários aspectos em conjunto, estaremos diante de um Prova/Teste.

Através de pesquisas longas e de profundidade, é possível desenvolver testes e provas em que as respostas não sejam verbais, exijam opção por figuras, ou confecção de desenhos, ou organização de peças ou figuras, enfim, atos cuja conexão com o conceito ou construto a medir não sejam imediatamente transparentes ao respondente. A expressão numérica do conceito corresponderá, então, ao score final apurado. O ponto alto desse tipo de quantificação é a forte independência da subjetividade pessoal, tanto do respondente como do agente. O aspecto controverso é ter-se certeza de que se está efetivamente medindo o conceito presumido. Se bem que tal preocupação exista em todas as situações de quantificação que não envolvem uma unidade de medida, nessa ela é ainda mais difícil de afastar, já que a conexão item-conceito é algo implícito, não diretamente visível. O caminho para se atingir o ajuste exige alterar e aperfeiçoar várias vezes tanto o conceito como o instrumento ou a prova, até que cheguem a algo geralmente aceito pela comunidade de pesquisadores. Em alguns momentos desse percurso, abre-se mão do enunciado explícito do conceito, como originalmente formulado, e se admite: “tal conceito é o que se mede através do instrumento tal” ou “pela escala de Fulano”.

Provas e testes têm também dois importantes espaços de aplicação. O de associar um valor numérico a um conceito ou construto é um deles, mas o segundo não é, como nos casos de item-escala, a coleta de opinião. É definir o grau de domínio do respondente acerca de determinado conteúdo, que encontra sua aplicação típica em provas de seleção (como concursos públicos) e de classificação (como os vestibulares), exames de suficiência (como os da OAB – Ordem dos Advogados do Brasil) e disciplinas curriculares em todos os níveis de escolaridade.

### 3.3 Um convite à quantificação

Nessa primeira parte deste artigo, procurou-se chamar a atenção para a importância e a utilidade da quantificação responsável e apresentaram-se as maneiras como isso pode ser concretizado. Tem-se, então, uma maleta de ferramentas. Pode ser interessante discutir rapidamente quando abri-la e que ferramenta usar.

A pesquisa, qualquer que seja, sempre se dá sobre uma base qualitativa. Ou algum fenômeno ou processo novo chama a atenção de alguém, ou algum resultado numérico precisa ser mais bem investigado – mas consciente de que *esse resultado exprime apenas um grau de conhecimento limitado acerca de um conceito que precisa estar qualitativamente formulado com clareza e rigor*, dentro de um contexto de processo e fenômeno cuja descrição completa é certamente bem mais verbal que numérica.

Sempre que viável, deve-se tentar medir através de contagem de repetições de um mesmo elemento básico, a *unidade*. Em alguns casos, a unidade é natural: quantos artigos o docente publicou em três anos? Qual o faturamento bruto em reais no mês de março? Quantos colaboradores tem a empresa? Em outros, pode-se criá-la, como fazem os físicos: Quantos milímetros (ou metros ou quilômetros) cobrem essa distância? Quantos segundos equivalem a esse intervalo de tempo? Essas são as quantificações mais fortes e confiáveis, o envolvimento de decisões subjetivas é mínimo, e elas podem, em muitos casos, ser realizadas de modo completamente automatizado.

Porém muitas importantes características de fenômenos e processos não permitem essa abordagem, não têm essa unidade, nem natural, nem construída. Por exemplo, quanta informação contém um artigo? Quão satisfeitos estão nossos clientes? Qual o risco social de um adolescente que mora numa favela? Quanto você ama seu parceiro ou parceira? Qual a qualidade de um curso de graduação (ou de pós-graduação)? E sua inteligência, quanto vale? São respostas que, eventualmente, se conseguem no nível dos indicadores, mediante o emprego de um instrumento ou de uma quantificação simples item-escala.

Sob o ponto de vista científico, as quantificações menos confiáveis são as que se constroem a partir de declarações pessoais, e mais ainda nos casos em que a resposta é determinada com base apenas em juízos subjetivos, exclusivamente internos ao respondente e, portanto, inacessíveis ao pesquisador. Por exemplo, considere responder de *fato* um “Como vai você?”. A resposta, seja verbal, seja em relação a uma escala explícita, é determinada em função de um conjunto maior ou menor de percepções internas de quem responde, condições essas raramente estáveis em termos de qualidade e de quantidade. Em instantes, a hierarquia e o número dessas condições mudam, e a resposta passa a ser outra: *Pensando bem, não é isso não... é o oposto!* A repetibilidade da resposta é incerta, tanto quando a questão é proposta ao mesmo indivíduo como quando formulada a vários deles. O valor científico, se algum, é muito pequeno.

Um passo à frente em confiabilidade e validade para pesquisa estão os casos em que a temática envolvida é bastante específica e o respondente é um especialista treinado, por exemplo, nas avaliações escolares. É lícito ter expectativa de que os resultados da avaliação sejam convergentes quando ela é repetida para o mesmo respondente, sejam as repetições efetuadas pelo mesmo avaliador, sejam por outro, igualmente preparado.

Novo progresso se obtém quando a quantificação não se fundamenta em declaração de pessoas, mas é feita mediante observação direta de uma situação ou contexto por um especialista. Exemplificando, é o caso das avaliações *in loco* de cursos de graduação e de pós-graduação pelos órgãos reguladores. A justa expectativa é que as notas ou conceitos atribuídos sensivelmente independam das pessoas responsáveis pelo processo avaliatório.

## 4 CLASSIFICAÇÕES DE INDICADORES

Tal como quantificar, classificar é também inerente ao ser humano. Identificar características comuns e diferenciadas e agrupar objetos de acordo com elas é um exercício por vezes suficientemente sedutor para levar pessoas a gastar muito dinheiro para montar coleções. Mas meramente separar e agrupar não é a finalidade das classificações científicas. Nelas, o propósito é utilizar a classificação para propósitos decisórios, para fazer escolhas, para identificar lacunas, para conhecer limitações e virtudes de algo, para definir oportunidades de utilização. A classificação científica não termina nela mesma, ela precisa ser



consequente, prestar-se a alguma finalidade.

Isso se aplica integralmente aos indicadores. Eles podem ser classificados de acordo com diversos critérios, todos relacionados a alguma característica que faz maior a sua compreensão ou permite obtê-los ou utilizá-los mais adequadamente. É o que passamos a discutir.

#### 4.1 Indicadores quanto à forma de determinação

**Diretos:** apuram-se a partir de um único conjunto de dados, todos de igual natureza, obtidos mediante repetições de um mesmo procedimento de observação do fenômeno de interesse, eventualmente empregando-se algum instrumento. A obtenção do valor final pode envolver algum tratamento matemático ou estatístico, mas sem qualquer participação de outros indicadores ou medidas.

**Indiretos:** são calculados a partir de *dois ou mais* indicadores diretos, tanto os obtidos por observação do fenômeno de interesse como aqueles eventualmente provenientes de outras fontes. O resultado exprime *um conceito distinto* de qualquer um dos participantes do cálculo. Por exemplo, *distância* quantifica a ideia de longe e perto, e *tempo*, a noção de demora. Já o quociente de ambos corresponde a *velocidade*, associada a algo rápido ou lento.

**Compostos (ou híbridos):** são também calculados a partir de outros, quase que exclusivamente mediante adição. Os participantes do cálculo correspondem a conceitos diferentes, o que, em princípio, torna a soma inaceitável sob um ponto de vista estritamente metrológico. No entanto, o interesse de quem computa o indicador não está no significado integral de cada conceito, mas em alguma característica que, esta sim, deve estar necessariamente presente em todos. Em outras palavras, para gerar um indicador híbrido, adicionam-se indicadores distintos no particular, mas similares num contexto geral, mais amplo. Por exemplo, pensando em um programa de pós-graduação, os números de artigos publicados, de teses defendidas, de dissertações aprovadas e de publicações em anais de eventos são indicadores todos distintos, mas compartilham a característica *produção científica*, e podem ser adicionados para caracterizá-la.

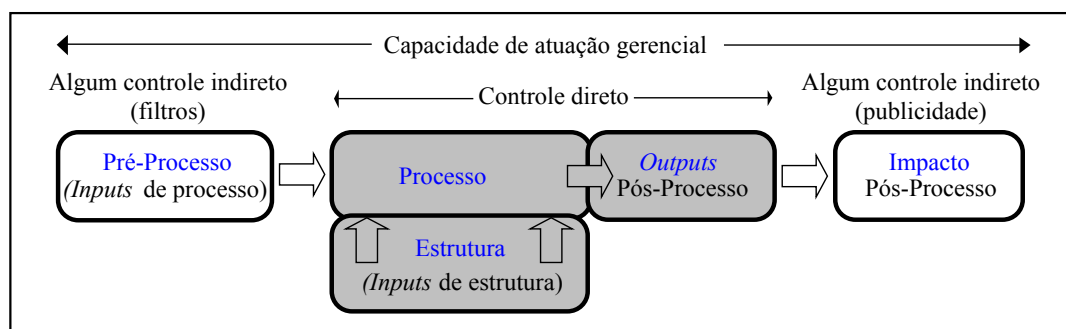
Note-se que, relativamente a um indicador de determinação indireta, um híbrido não está associado a um novo conceito, mas a algo que lhe é inerente, já existente em todos os indicadores que irão entrar em sua composição.

**Proxys:** podem ocorrer situações em que não exista qualquer método ou possibilidade de observar a característica de que se necessita, ou que seja demasiadamente complexo fazê-lo. Uma alternativa é, então, adotar um indicador *proxy*, um outro indicador que se supõe presente suficiente correlação com o conceito em que se esteja interessado para servir como sua respectiva quantificação. Um exemplo é usar o produto interno bruto *per capita* como *proxy* para o padrão ou a qualidade de vida de uma população.

#### 4.2 Indicadores quanto à parte do processo a que dizem respeito

A primeira seção deste artigo apresenta os indicadores como aliados imprescindíveis na construção do conhecimento, que conceitua como a capacidade de fazer o Universo evoluir no sentido desejado e definido *a priori*. Em outras palavras, inibir os fenômenos e processos desfavoráveis e promover os favoráveis. Essa visão nos oferece um segundo critério para classificar indicadores: quanto à etapa do processo a que se referem e quanto à capacidade de atuação gerencial, no sentido de controlar as condições que definem seu valor. Temos aí os cinco tipos a seguir, ilustrados na Figura 1.

Figura 1- etapas de um fenômeno passíveis de caracterização mediante indicadores



Fonte: elaborado pelo autor (2014)

De *estrutura (ou inputs de estrutura)*: dizem respeito a aspectos do processo que, em princípio, poderiam variar, mas não há a expectativa de que isso aconteça durante o transcorrer do fenômeno. Têm

as características de *parâmetros de configuração*, ou seja, seus valores são estáveis no tempo, relativamente à dinâmica do processo, mas podem ser modificados (por exemplo, aquisição de novas máquinas por uma empresa, mudança de instalações físicas, ampliação do quadro de colaboradores em um setor e similares). As mudanças tendem a ter caráter definitivo, acontecem em decorrência direta da ação gerencial e os intervalos de tempo entre elas são em geral bastante longos.

De *pré-processo* (ou *inputs de processo*): referem-se a quaisquer características relevantes existentes antes do processo (antes que ocorra o ingresso na estrutura e que o processo propriamente dito se inicie) que possam vir a influenciar o processo ou seu desfecho. A gestão do processo pode ter algum controle indireto sobre o que espera dessa fase, *filtrando* (aceitando ou rejeitando) ingressos que não atendam determinadas especificações, mas não pode, em princípio, modificá-las. Elementos ou ingressos não passíveis desse tipo de controle prévio serão necessariamente incorporados ao processo no estado em que se encontram (por exemplo, indicadores de risco do contexto, como instabilidade social, alterações na política econômica e similares).

De *processo*: refletem informações relevantes sobre o andamento do processo, permitindo monitorá-lo para fins de auxílio à decisão e acompanhamento gerencial; são dinâmicos, ágeis no curto prazo, modificando-se em tempo real (ou com bom *sincronismo*) relativamente às mudanças da informação que traduzem (o que caracteriza esses indicadores como *variáveis*, e não *parâmetros*). Possibilitam a definição de metas intermediárias ao longo do processo.

*Outputs*: dizem respeito à fase *pós-processo*. Refletem características e especificações dos produtos finais e são fortemente controláveis pela gestão.

De *impacto*: são também *pós-processo*, mas diferenciam-se porque, sobre eles, a gestão tem uma influência *no máximo* distante e indireta (por exemplo, através de publicidade). São criados para atender a uma *demand gerencial*, quantificando, para o gestor, a eventual repercussão dos produtos e serviços que o processo entrega à sociedade. Frequentemente, indicadores de impacto suprem igualmente bem a *demand da comunicação*, em que seu alvo é o público, e não o gestor.

Quando um fenômeno ou processo é caracterizado ou avaliado por um conjunto de indicadores, é interessante e útil verificar como eles se distribuem pelas cinco categorias recém-descritas. Isso permite detectar excesso ou falta de ênfase relativamente a alguma delas, o que pode tanto ser decorrência de equívoco ou omissão na construção do conjunto, como expressão deliberada de preferência política em uma avaliação.

#### 4.3 Indicadores quanto ao potencial de emprego gerencial

*Causais* (ou *de ação*): dão informações relacionadas somente a causas que afetam (positiva ou negativamente) o andamento do processo ou seus resultados. Quando empregados com finalidade gerencial, propiciam soluções que induzem à melhoria ou afastam o problema definitivamente. Por exemplo, o número de focos do mosquito da dengue num determinado bairro (ter em mente que a relação causal em questão pode ser estocástica ou determinística).

De *efeito* (ou *de controle*): informam apenas se o processo ou seus resultados estão adequados ou não, mas não possibilitam a determinação de uma causa única, clara, para o contexto observado. Podem ou não dar ensejo a ações gerenciais, mas, no caso positivo, as soluções que propiciam têm um efeito temporário apenas sobre a característica específica que o indicador expressa, não eliminando as causas que o fazem inadequado – isto é, ataca-se o sintoma, não o que o origina e, terminada a ação, o efeito retorna. A temperatura corporal é um bom exemplo de *indicador de efeito*. Se estiver acima de 37°C, implica que algo não está bem, porém nada permite concluir acerca da causa da temperatura elevada. Por outro lado, induz a uma ação gerencial: será necessário baixar a temperatura, para que a pessoa não corra risco de convulsão e morte.

De *efeito/causa*: combinam as duas características anteriores, informando um resultado cuja causa é única e bem determinada, ou seja, dizem respeito a um encadeamento causa-efeito determinístico. Assim, atendem simultaneamente os requisitos para controle/avaliação e para ação/intervenção.

#### 4.4 Indicadores quanto à amplitude da informação que abarcam

*Singulares* (*simples/básicos/primários*): contêm uma informação única, clara, bem definida relativa ao processo, podendo ser de determinação direta ou indireta. A temperatura corporal é, também aqui, um bom exemplo. Quando forem causais ou de efeito/causa, serão os mais úteis e eficazes para a intervenção gerencial.

*Mesossintéticos*: são do tipo composto, determinados a partir de vários indicadores, a maioria deles singulares, os quais dizem respeito às várias características (consideradas relevantes) de um sistema ou processo. Visam a obter a respectiva apreciação global, expressa através de um valor único. Indicadores desse tipo são úteis para avaliações de setores ou partes de sistemas ou processos mais amplos, como os

indicadores corpo docente ou infraestrutura numa avaliação de instituição de ensino superior. São capazes de informar se algo vai bem ou não, mas dificilmente servem para decisões operacionais de gestão.

*Macrossintéticos*: são também compostos, mas predominantemente obtidos a partir de indicadores mesossintéticos, avaliando o sistema ou processo mais amplo a que estes últimos dizem respeito. Por exemplo, os conceitos dos cursos de Pós-Graduação (CAPES) e o produto interno bruto de um estado ou país. Como os mesossintéticos, informam se algo vai bem ou não, mas dificilmente servem para decisões operacionais de gestão.

## 5 FUNÇÕES DOS INDICADORES

Existem ao menos três importantes demandas com respeito às quais o emprego de indicadores é especialmente útil e eficaz: a gerencial, a de avaliação e a de comunicação.

*Demanda Gerencial*: uma empresa, um projeto, uma aeronave são sistemas em processo para os quais se almeja um desfecho positivo. Isso exige conhecimento, para conduzir o Universo no sentido desejado, e informações, para as tomadas de decisão adequadas. Relembrando Lorde Kelvin, o conhecimento quantificado é muito rico e objetivo. Então a concepção, a coleta, o entendimento e a correta interpretação dos indicadores relativos a um fenômeno ou processo são vitais para a sua gestão adequada. Considere, por exemplo, o pouso por instrumentos de uma aeronave. As decisões da tripulação baseiam-se fortemente nos indicadores fornecidos pelos equipamentos de bordo e de terra.

Um indicador atende adequadamente a demanda gerencial quando o gestor sabe exatamente o que é preciso fazer para ajustar o seu valor, sem a necessidade de retornar ao sistema, em busca de informações adicionais, ou de recorrer a outros indicadores, mais simples e diretos.

Sob o ponto de vista da gestão empresarial, indicadores não apenas municiam o gestor com informações para conduzir a dinâmica do processo. Eles também são usados para fixação de *metas* progressivas (ou *valores de referência internos*), a serem atingidas no decorrer do processo (aqui, a fonte é o gestor, pessoa ou pesquisador no controle). Bem utilizadas, as metas podem ser importantes fatores de motivação para uma equipe. Mas podem ter um efeito contrário se forem excessivamente duras e difíceis: desmotivação, estresse, tensão, desavenças. Notar, porém, que o indicador é sempre inocente: o que pode ser perverso, se mal empregado, são as metas.

Uma segunda aplicação de indicadores com vocação gerencial é o estabelecimento de *padrões* normativos ou de excelência (ou *valores de referência externos*), os quais o processo precisa atender. Diversamente das metas, os padrões não são fixados pelo gestor, mas por ou em função de alguma influência externa – órgão regulador ou resultados de um concorrente.

*Valores de referência* podem ser estabelecidos na forma de *piso* – um valor mínimo que precisa ser atingido – ou de *teto* – um valor máximo que, quando alcançado, esgota os benefícios que oferece ao processo ou, pior, começa a causar efeitos deletérios.

*Demanda de Avaliação*: nesse domínio, os indicadores são planejados e utilizados para colher informações sobre o andamento do processo e a qualidade do seu desfecho. Objetivos são, entre outros, verificar atendimento a normas, certificação, concessão de financiamento, atribuição de conceitos e prêmios. Indicadores de avaliação não precisam ser concebidos com a preocupação direta de aperfeiçoar o processo, e talvez alguns deles não possam mesmo oferecer essa possibilidade. Mas induzir a melhorias no que está sendo avaliado – ser construtivo, jamais punitivo – constitui-se em uma qualidade importante e, até, imprescindível de boas avaliações. Ao criar, analisar ou criticar um indicador de avaliação, deve-se sempre considerar o seu potencial de aperfeiçoar o processo a que se refere.

Avaliações de caráter amplo envolvem todos os tipos de indicadores, dos singulares aos macrossintéticos, e, direta ou indiretamente, acabam estabelecendo padrões para vários deles. É uma importante habilidade gerencial transformar os indicadores de avaliação em outros, mais simples, específicos, que possam ser diretamente operacionalizados e acompanhados em tempo real, a fim de direcionar ações e intervenções no sentido do atendimento dos padrões no longo prazo.

*Demanda da Comunicação*: finalmente, existem indicadores destinados prioritariamente a comunicar bem. São, em geral, macrossintéticos relativamente ao seu contexto, e têm que fazer sentido e impressionar (positiva ou negativamente) as pessoas a que se destinam. Em termos de *comunicação especializada* pode-se pensar em um gerente com um grande projeto em andamento que precisa convencer a diretoria da empresa a ampliar sua verba, numa reunião em que poderá apresentar cerca de meia dúzia de informações. Indicadores bem selecionados ou especificamente criados fazem a diferença aí. Ampliando agente e público, considere as empresas de capital aberto que precisam manter uma imagem positiva entre os analistas e investidores: que indicadores incluir no extrato do balanço e em seus *press releases*? Veja-se, por exemplo, a questão ambiental, que não apresentava a atual relevância na primeira metade do século XX. Por fim, indo efetivamente para o macroscópico (*comunicação não-especializada*), governos têm que impressionar

a população com seus feitos e realizações. O PIB é um indicador de bom impacto e comunicação? O número de atendimentos e internações propiciados pelo Sistema Único de Saúde? A queda nos casos registrados de dengue? O enriquecimento médio anual dos políticos do “outro” partido? O montante total despendido pelos turistas estrangeiros no país, durante um evento esportivo mundial? Em todas as situações, decidir quantos e quais indicadores selecionar exige muito discernimento: uma quantidade excessiva enfraquece a impressão que cada um causaria isoladamente, o que tende a prejudicar a efetividade da comunicação.

## 6 INDICADORES: GARANTIA DA QUALIDADE

### 6.1 A apresentação do indicador

Todo indicador deve possuir um *documento de apresentação*, constituído de cinco partes: *denominação*, *propósito*, *conceito*, *forma de apuração* e *metadados*.

A qualidade do indicador precisa se afirmar desde logo, na sua *denominação*, que já deve propiciar uma boa noção do seu *conceito* e de seu *propósito*. Por exemplo, *Taxa de ocupação* é um nome de indicador adequado para empresas aéreas, restaurantes e estabelecimentos de hospedagem, e é bastante transparente quanto ao propósito. Deixa a desejar, no entanto, quanto a refletir o conceito, que exigiria complemento temporal (*instantânea em “tal” momento* ou, então, *média diária* ou *mensal* ou *sazonal* etc.).

Outros dois componentes básicos da qualidade do indicador são o seu *propósito* e o seu *conceito*. O primeiro é a resposta a *para quê?* e tem que estar explícito e claro, ele estabelece a razão de ser do indicador: que sentido existe em apurá-lo, se não apresenta uma finalidade definida? O segundo responde a *o quê?*, qual a informação de interesse acerca do processo ou do sistema que esse indicador monitora.

O quarto requisito essencial é a *forma de apuração* (o *como?*), a qual deve ser fornecida de modo explícito e detalhado, a ponto de o resultado independar da pessoa que efetua a determinação. Ou seja: caso, em qualquer fase da apuração (da coleta de dados até o cálculo final), duas ou mais pessoas tenham acesso à respectiva situação de momento e, a partir daí, passem a seguir as etapas preconizadas sem mais contato entre si, é preciso que todas essas pessoas cheguem necessariamente ao mesmo resultado. Isso está longe de ser trivial, como pode ser exemplificado considerando-se o indicador taxa de ocupação de assentos.

Vamos pensar em uma situação grandemente simplificada: uma empresa aérea com três aeronaves, cada uma das quais realiza um vôo diário, e cujas capacidades são 60, 120 e 200 assentos. Num dado dia, os vôos registraram 48, 90 e 50 passageiros respectivamente, correspondentes, nessa ordem, a taxas de ocupação de 80,0%, 75,0% e 25,0%. Bem, como determinar a taxa de ocupação diária? Os totais de passageiros e de assentos ofertados são 188 e 380, o que daria uma taxa de 49,5%. Porém a média dos percentuais antes calculados é 60,0%. Qualquer dos dois cálculos pode ser adotado, mas a opção feita tem de constar obrigatoriamente no documento de apresentação.

A última parte da apresentação do indicador é a listagem dos *metadados* necessários à sua determinação, isto é, de todas as variáveis, indicadores e outras informações que precisam ser conhecidas para encontrar o seu valor. Retomando a recém descrita taxa de ocupação de assentos, teríamos como metadados: para cada vôo realizado, (i) a data, (ii) o número ou o horário, (iii) o número de assentos ocupados, (iv) o número de assentos ofertados, e (v) os percentuais de ocupação (somente se o cálculo final fosse a média desses percentuais).

### 6.2 Requisitos da qualidade

Trzesniak (1998) amplia e adapta um conjunto de critérios para a formulação de bons indicadores, a partir de Tironi et al. (1991). Na dissertação de Araújo (1999), dois mais são adicionados (*frequência e não-interferência*), e outros acréscimos estão sendo feitos neste texto (*coerência global e sincronismo*). O conjunto completo, apresentado nesta seção e na próxima, divide os requisitos em duas categorias: de *qualidade* e *desejáveis*.

A categoria de *qualidade* engloba as propriedades a seguir que, quanto mais bem forem atendidas pelo indicador, mais o farão útil e eficaz.

- *Coerência global*: o nome e a forma de apuração do indicador devem estar rigorosamente sintonizados com seu propósito e seu conceito, e esses últimos entre si.

- *Frequência*: o indicador deve retratar uma característica que chame a atenção do pesquisador por sua *frequência de aparecimento*, isto é, por ser percebido todas (ou quase todas) as vezes em que o fenômeno sob investigação se repetir. Caso ela esteja também envolvida em outros fenômenos, o indicador respectivo torna-se forte candidato a atender ao menos o requisito desejável *abrangência*, eventualmente até o *portabilidade*.

- *Univocidade*: o indicador deve retratar um aspecto único, claro e bem definido do processo. A

necessidade desse requisito é crítica para os indicadores gerenciais, mas pode ser amenizada nos de avaliação e de comunicação.

- *Relevância* ou, segundo Tironi et al. (1991), *seletividade*: o indicador deve retratar um aspecto importante, essencial, crítico do processo/sistema.
- *Gradação de intensidade*: o indicador deve variar (não permanecer *constante*) durante o intervalo de tempo típico de ocorrência do processo e no espaço dos processos de interesse.
- *Sincronismo (fidedignidade dinâmica ou rapidez de resposta)*: o indicador precisa reagir rapidamente às variações da característica que ele monitora no processo (quanto menor a demora entre a variação e a resposta/atualização de valor do indicador, melhor).
- *Não-interferência*: deve ser possível extrair as informações necessárias à determinação do indicador sem alterar a configuração e a evolução futura do processo.
- *Baixo custo de obtenção*: É muito favorável (talvez até decisivo) para o sucesso dos indicadores que eles possam ser obtidos com custo baixo (TIRONI et al., 1991), preferencialmente que sejam determinados a partir de dados já disponíveis acerca do processo em estudo (como os dados contábeis em uma empresa).
- *Padronização* ou, segundo Tironi et al. (1991), *estabilidade*: a geração do indicador deve basear-se em uma norma, um procedimento único, bem definido e estável no tempo.
- *Rastreabilidade*: os dados em que a obtenção do indicador é baseada, os cálculos efetuados e os nomes dos responsáveis pela apuração devem ser registrados e preservados (TIRONI et al., 1991).

### 6.3 Requisitos desejáveis

Já os requisitos desejáveis não são previsíveis ou esperados *a priori*, mas correspondem à possibilidade de se empregar o indicador em processos/sistemas diversos daqueles no âmbito dos quais ele foi inicialmente concebido, sem que sejam perdidos a sua validade, o seu significado, a sua fidedignidade relativamente à informação que retrata. Esse é um passo frequentemente subestimado, e indicadores migram de uma aplicação para outra de um modo quase irresponsável. *Toda e qualquer ampliação de escopo de um indicador tem de ser exaustivamente testada e validada.*

São características desejáveis:

- *abrangência ou amplitude*, ou *cobertura* segundo Tironi et al. (1991): amplia-se quando a nova aplicação corresponde a processos/sistemas *de natureza diferente*, porém *dentro da mesma (sub)área do conhecimento* na qual o indicador é usualmente empregado. Tem-se aumento da abrangência quando se constata que um indicador, concebido para hotéis, pode ser aplicado com igual sentido e eficácia a pousadas e spa's. Outra extensão seria poder usá-lo também para motéis e, ainda outra, para restaurantes. Um segundo exemplo é usar para pesquisadores (pessoas) um indicador de produtividade em pesquisa criado para avaliar universidades (instituições).
- *Transferibilidade (portabilidade)*: similar, porém mais forte, significa que o uso do indicador pode ser estendido com sucesso para o estudo de processos de outras (sub)áreas do conhecimento contíguas a alguma em que ele funciona adequadamente. Por exemplo, aplicar ao setor industrial um indicador criado para o de serviços.
- *A invariância* de escala está presente quando o indicador mantém sua validade e sua interpretação, mesmo que ao menos uma das *dimensões* relevantes do processo examinado seja acentuadamente diferente. Por exemplo: um indicador que tem bom desempenho na gestão de empresas de até 50 colaboradores continua fazendo sentido se esse número aumenta para 10000? Número de artigos publicados por docente, em uma instituição com 80 professores, tem o mesmo significado e a mesma interpretação em outra, com 5000?

## 7 CONSIDERAÇÕES FINAIS: RESPONDENDO A DESAFIOS COM DESAFIOS

O presente artigo coloca, para quem o estuda, uma série de desafios: (i) aproxime-se da quantificação e descubra como ela pode ser preciosa; (ii) entenda e tenha sempre em mente a possibilidade de estar diante de um processo estocástico; (iii) conheça e esteja atento para a confiabilidade e validade de medidas e indicadores; e (iv) conheça bem os indicadores, sabendo classificá-los e criticá-los, a partir da identificação de suas características.

A causa para a estruturação do texto em desafios – a propósito, a mesma que o faz ter um número tão pequeno de referências – é que ele foi construído em resposta a desafios, e não com base em muitos outros textos. Ao elaborarem propostas de avaliação de indicadores existentes a partir da categorização em *de qualidade* e *desejáveis*, e ao convidarem o autor a participar desse exercício, dois pesquisadores e uma mestranda do Programa de Ciência da Informação da Universidade Federal de Pernambuco ofereceram inúmeras oportunidades de pensar e repensar o tema, de estender e ampliar visões, de adaptar e criar.

O relatório de Presser (2013), o projeto de Santos (2013) e a dissertação de Silva (2013) guardam essa memória. Esses trabalhos, assim como muitas das referências neles citadas, com destaque para Araújo (1999) e Bertolin (2011), sem dúvida contribuíram significativamente para que este artigo viesse a ser elaborado. No entanto, as contribuições não são do tipo pontual, que se limitam a um ou outro parágrafo e podem ser convencionalmente indicadas ao longo do texto. Mais que isso, esses autores chamaram a atenção para cenários e alternativas para a concepção, apuração e emprego de indicadores – os desafios – que o presente texto buscou contemplar. A todos eles, o mérito da provocação. Ao autor, a responsabilidade pela resposta.

## QUANTITATIVE INDICATORS: HOW TO OBTAIN, EVALUATE, CRITICIZE AND IMPROVE

### ABSTRACT

Quantifying is inherent to human beings, and it is undeniable that bringing this practice into the construction of knowledge was a decisive factor for the great success of the (hard) sciences and engineering in the nineteenth century. Thus, the transformation of information into numbers to understand and control processes was transferred to the Humanities and Social Sciences... but success was not as great as expected, leading many researchers of these areas to reject quantification at all. Although this attitude is understandable, we argue that it does not result from quantification itself, but that, along with the latter, subliminally came two very questionable and insufficiently discussed assumptions: that, as in Physics, in the Humanities and Social Sciences, (i) the phenomena are predominantly deterministic, and (ii) there is the expectation of a clear and well defined “true value” for the quantified information. Actually, quantification in these no-exact areas becomes much more useful and effective if one recognizes that: (i) the phenomena are better understood if approached as stochastic, (ii) there is no expectation of a “true value” for the quantified information; variables and parameters are predominantly distributions (numerical intervals), and (iii) the ways to quantify, although undoubtedly fully valid, are less objective. In this article, we characterize each of these contexts, in order to discriminate them clearly and definitely. We differentiate between measurements and indicators, and focus on the latter, classifying them according to various features, functions, and applications, and concluding with a list of quality requirements that they must obey to maximize their effectiveness, usefulness and accuracy.

**Keywords:** Indicators. Measurements. Stochastic processes. Quantification. Modeling.

### REFERÊNCIAS

ARAÚJO, D. G. **Indicadores quantitativos para os processos de ensino-aprendizagem:** caracterização da repetência e da evasão. 1999. Dissertação (Mestrado) - Escola Federal de Engenharia de Itajubá, Itajubá, 1999.

BERTOLIN, J. C. G. Uma proposta de indicadores de desempenho para a educação superior brasileira. **Est. Aval. Educ.**, São Paulo, v. 22, n. 50, p. 471-490, set./dez. 2011.

PRESSER, N. H. **Reflexões que antecedem a construção de um sistema de informação de gestão acadêmica:** estudo de caso em programas de pós-graduação (Relatório técnico, processo CNPq 401896/2011-0). Recife: Departamento de Ciência da Informação/UFPE, 2013.

RESNICK, Robert; HALLIDAY, David. **Física I.** 2. ed. Rio de Janeiro, RJ: Livros Técnicos e Científicos, 1973.

SANTOS, R. N. M. dos. **Formulação de diretrizes para concepção e implantação de um observatório na UFPE** (Projeto de pesquisa). Recife: Departamento de Ciência da Informação/UFPE, 2013.

SILVA, E. M. da. **Aspectos Contextuais e Metodológicos da Avaliação da Qualidade da Educação Superior Brasileira:** uma análise dos indicadores utilizados pelos órgãos reguladores nos processos de avaliação dos cursos de graduação. 2013. 129 f. Dissertação (Mestrado Profissional em Gestão Pública para o Desenvolvimento do Nordeste) – Universidade Federal de Pernambuco, Recife, 2013.

TIRONI, L. F. et al. **Critérios para geração de indicadores da qualidade e produtividade no serviço público** (Texto para discussão #238). Brasília-DF: Instituto de Pesquisas Econômicas Aplicadas-IPEA, 1991.

TRZESNIAK, P. Indicadores quantitativos: reflexões que antecedem seu estabelecimento. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 159-164, maio/ago. 1998.