

Será que a substituição dessas formas tão imperfeitas de mensuração por classificações qualitativas é a solução?

Creio que a chamada avaliação qualitativa tem também suas próprias dificuldades e inúmeros desafios.

Para ficar apenas no âmbito da avaliação de pesquisadores individuais, embora pareça tentadora a ideia de que o próprio autor determine quais são as suas produções mais significativas e as ofereça à avaliação por pares, na prática não parece tão simples.

Todos os editores científicos e os coordenadores de processos de avaliação de projetos para o fomento têm enfrentado enormes dificuldades com a revisão por pares. Cada vez torna-se mais difícil conseguir avaliadores e avaliações bem feitas que possam ser realmente úteis para as comissões encarregadas da aprovação ou priorização das propostas.

Quando as avaliações envolvem projetos de diferentes áreas e subáreas como é possível hierarquizar as propostas? Como avaliar projetos de áreas ou subáreas que se encontram em diferentes estágios de desenvolvimento? Será que uma produção considerada de excelência em uma determinada área pode ser comparada com outra, avaliada em uma área distinta?

Em campos como a Saúde Coletiva, com uma enormidade de objetos, abordagens, correntes teóricas, temáticas, posicionamentos ideológicos, parece extremamente desafiador a proposição de avaliações classificatórias.

Como garantir que os comitês julgadores sejam adequadamente constituídos para dar conta dessa diversidade sem cometer grandes injustiças?

Enfim, todas essas reflexões foram aqui apresentadas apenas no intuito de problematizar uma questão que, me parece, está muito longe de ser adequadamente equacionada.

Nós todos que nos interessamos pela ciência como objeto de investigação temos muito com o que nos ocuparmos. Sem contar com o enorme desafio que significa tentar avaliar o real impacto da ciência e de seus produtos na melhoria das condições concretas de existência da humanidade.

<http://dx.doi.org/10.1590/0102-311XCO020913>

Paulo Vaz

*Escola de Comunicação,
Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.
paulorgvaz@uol.com.br*

Por um novo jogo

Como a pesquisa no Brasil depende quase que totalmente de financiamento público, a distribuição de verbas tende a utilizar formas ordenadas de avaliação. É preciso responder prévia e estruturalmente, por meio da adoção de procedimentos, a possíveis denúncias de injustiça nos julgamentos das agências de fomento.

Visando a assegurar legitimidade e eficiência na alocação de recursos escassos, nos últimos 20 anos, dois tipos maiores de indicadores quantitativos paulatinamente dominaram as avaliações: um calcado na hierarquização dos veículos nos quais as pesquisas são publicadas, principalmente a classificação de revistas científicas, e o outro na repercussão do trabalho do pesquisador, mensurada pelas citações que recebeu.

O artigo de Keneth Rochel de Camargo Jr. propõe que é hora de avaliar o uso praticamente absoluto dessas duas métricas nas avaliações das ciências no Brasil. A ênfase do texto recai na exposição e crítica de algumas das razões que nos levam a confiar nesses indicadores. Dentre as razões, podem ser destacadas a crença nos números como garantidor da objetividade, a facilidade de construção de consenso e a possibilidade de os juízes se eximirem do risco de julgar, isto é, de se eximirem da necessidade de justificar suas escolhas perante seus pares.

O artigo também aponta brevemente as consequências negativas. As avaliações, hoje, quando se limitam ao uso dessas métricas, tendem a não ser capazes de apreender a qualidade e favorecem a consolidação de um ambiente marcado pelo privilégio da comunicação em detrimento da informação, ou seja, pelo privilégio de publicar em detrimento de realizar pesquisas de qualidade. Esse “produtivismo” também tende a médio prazo a esgarçar as comunidades científicas, pois o volume de publicações e a necessidade de publicar podem fazer com que os pesquisadores especializem cada vez mais suas leituras e tenham muita dificuldade para saber o que é relevante e inovador devido à deterioração da relação entre sinal e ruído.

Este comentário irá se deter sobre o produtivismo, essa consequência negativa global. Não é difícil compreender por que esses índices foram criados. Ter seu trabalho publicado nas melhores revistas de uma determinada área significaria, em princípio, ter enfrentado uma competição mais dura e ter sido avaliado com mais rigor por pares do que seria o caso numa revista menos importante. E ter sua pesquisa citada é o modo

1. Barreto ML, Aragão E, Souza LEPF, Santana T, Barata RB. Diferenças entre as medidas do índice-h geradas em distintas fontes bibliográficas e engenhos de busca. *Rev Saúde Pública*; no prelo.

de indicar seu valor, pois serviria de apoio e inspiração para outras pesquisas.

Há um elemento adicional nessas métricas que é relevante. Seu uso implica a multiplicação dos juízes que decidem a qualidade de um pesquisador. Ao invés de ter seu trabalho reconhecido apenas pelos membros de um dado comitê, ele foi avaliado por outros pesquisadores, seja para ser publicado, seja por ter sido citado.

Até aqui, portanto, nada teria sido melhor do que a construção desses índices. Cabe, porém, um alerta cético. Toda vez que formas de hierarquizar o mérito são objetivadas por meio de regras, o que se produz é um outro modo de indivíduos e grupos lutarem entre si. Um novo procedimento não traz a paz; cria, sim, novos modos de desdobrar conflitos. Consideremos a disputa entre psicanálise e ciências cognitivas que afeta a psicologia e a psiquiatria. O uso de critérios quantitativos tende a favorecer os praticantes das ciências cognitivas, que publicam mais em revistas do que em livros e que se pensam, acima de tudo, como cientistas e não como intelectuais a questionar a sociedade. Esse favorecimento não é destino; os psicanalistas podem criar revistas próprias, lutar para que livros sejam considerados e propor a necessidade de valorizar o reconhecimento do público leigo.

A existência de regras também produz o esforço de manipulá-las. Quando os juízos eram preferencialmente qualitativos, manipulava-se por se aproximar dos avaliadores. Em relação às regras atuais, já são bastante criticados o clube da citação recíproca e a multiplicação indevida da coautoria.

Em suma, quando a distribuição do valor dos indivíduos decorre de um julgamento, haverá questionamento tanto de quem julga, quanto dos critérios aplicados. Haverá, também, esforços para distorcer as regras a favor de grupos ou indivíduos. Haverá, por fim, efeitos gerados pela aceitação consentida. A maior parte dos pesquisadores não manipula, mas joga segundo as regras que estimulam a produção e, assim, publicam cada vez mais. Perdemos de vista aquilo que as métricas visavam medir: a qualidade da pesquisa. A publicação-salame não é uma jogada faltosa, como a falsa coautoria; é, sim, um modo de jogar tornado possível pelas novas regras. Quem o adota só pensa em vencer, não se importando mais com a singularidade do jogo que está jogando.

Os problemas com esses índices não se limitam, portanto, à incapacidade local de reconhecer a qualidade ou ao privilégio de grupos e visões de ciência. Em termos globais, o horizonte negativo desdobrado pela forma de julgar apoiada apenas em métricas de publicação teria como

figuras maiores o já mencionado esgarçamento das comunidades de pesquisa pela deterioração da relação entre sinal e ruído, e a possibilidade de as práticas científicas não terem mais como eixo normativo a descoberta da verdade.

Diante desse horizonte, propomos dois tipos de questão ao texto. Em primeiro lugar, e tendo em vista a necessidade de reduzir os incentivos ao produtivismo inerentes à forma atual de avaliação, não é necessário estabelecer uma diferença entre avaliação de revista e citação? Como aumentou o número de revistas e de edições, a classificação hierárquica não foi obstáculo ao produtivismo. Em alguns casos, como quando se usa os estratos Qualis da Capes, ocorreu, ao contrário, aumento das oportunidades de publicação. Já a citação incentivaria o pesquisador a privilegiar a relevância. Se o que importa é publicar o que servirá de apoio a outras pesquisas, não há por que dividir uma pesquisa em vários trabalhos ou publicar só por publicar. Mais do que abandonar essa métrica, talvez o que caiba fazer seja sofisticá-la para evitar as imperfeições realçadas no texto.

Em segundo lugar, explicitamente, o texto se opõe apenas ao uso exclusivo de critérios quantitativos. Mas não insiste, como creio ser necessário, na necessidade de conjugar o qualitativo com as métricas quantitativas. Imediatamente, basta imaginar o que aconteceria em campos de conhecimento conflagrados se a avaliação da pesquisa fosse reduzida ao juízo qualitativo de membros de um comitê. Dependendo da composição, seria fácil excluir o que é divergente em relação ao ponto de vista dos avaliadores. Mesmo em áreas sem conflito aberto, o juízo qualitativo de cientistas experientes – a alternativa ao uso de critérios quantitativo – estará marcado por vieses afetivos e os derivados de visões de mundo. As métricas quantitativas, por multiplicarem os avaliadores, podem ser instrumentais na redução do arbítrio inerente ao juízo qualitativo. Os comitês terão de encontrar justificativas consistentes para recusarem financiamento a pesquisadores que publicam em revistas reconhecidas e que são bastante citados.

Por fim, um comentário sobre a sugestão presente no texto de reduzir o número de publicações a serem consideradas de modo a tornar exequível uma avaliação qualitativa. O que se visa não é apenas a criação de condições que possibilitem o julgamento qualitativo; também são reduzidos os incentivos a publicar muito, pois não precisa e não adianta. Em acréscimo e para além da busca do prestígio e da vitória, talvez o mistério volte então a ser força a mover o jogo científico.