

Avaliação

João Lucas Marques Barbosa

Quero agradecer aos organizadores do evento, em particular a SBF e mais particularmente ao Alaor, o convite para que eu pudesse participar deste evento.

Antes de tudo gostaria de registrar que o processo de avaliação pelos pares, utilizado no mundo inteiro como forma de julgamento da qualidade, é o mais adequado para a tomada de decisões para concessão de recursos para pesquisa, ou para a concessão de prêmios por mérito científico. É fundamental que se garanta a continuidade e o aprimoramento de tal sistema em contraposição aos pedidos políticos ou com base em amizade ou parentesco. Deve-se ter particularmente cuidado para que não se abra mão do julgamento pelos pares mesmo quando o governo insiste em que a concessão de recursos siga algum plano de desenvolvimento que considera fundamental para o País. Sigam-se os planos, mas mantenham-se os julgamentos do mérito pelos pares.

- A) Os procedimentos de avaliação pelo mérito científico, no caso brasileiro situam-se em varias instituições onde atendem a diferentes objetivos. As principais instituições são o CNPq, a CAPES, as FAPs, e as Universidades. A avaliação recai sobre pedidos de recursos para pesquisa, solicitações de bolsas, avaliação dos cursos de pós-graduação, etc.
- B) Embora todos os julgamentos tenham por base a avaliação do mérito científico (do pesquisador, do curso, do grupo de pesquisa, etc.) seus objetivos são bem distintos. Por exemplo:
 - a. No julgamento dos cursos de pós-graduação o objetivo é o de fazer crescer de forma organizada e incentivar a qualidade dos cursos de PG de forma a responder às necessidades do País. Evidentemente a principal consequência de tal julgamento ocorre no âmbito da própria CAPES, que utiliza o resultado da avaliação para a concretização de uma política de desenvolvimento científico e tecnológico para o País.
 - b. No julgamento dos pedidos de auxilio por qualquer das instituições, o objetivo é o de tomar uma decisão pontual de conceder ou não recursos para um projeto apresentado dentro de um Edital.
- C) Embora os processos de avaliação ocorram dentro das instituições do sistema, a avaliação em si é realizada pelos próprios pesquisadores e, conseqüentemente depende da concepção que cada indivíduo faz do que é um bom projeto, um bom curso de pós-graduação e, no background de tais decisões, o que é fazer ciência em cada área específica. É importante observar que, quando os resultados do julgamento de um edital, ou do processo de avaliação dos cursos, são divulgados, eles passam por uma crítica da própria comunidade de cada área que de certo modo avalia se os resultados então em consonância com o que ela

espera com base em um conhecimento latente da qualidade de cada programa ou de cada pesquisador. Quando há discordância, o responsável pela área na instituição recebe informalmente, e às vezes formalmente, reclamações argumentando contra a decisão. Nos casos que acompanhei este tipo de patrulhamento das decisões de comitê pela própria comunidade levou ao reconhecimento de erros crassos e a sua correção. Este acompanhamento feito pela comunidade é positivo para que o sistema de avaliação pelos pares funcione de forma adequada. Dentro desta mesma ordem de idéias, vale ressaltar que sendo as avaliações realizadas por subgrupos da comunidade, qualquer mudança significativa no processo de avaliação não ocorrerá sem uma mudança de concepção na própria comunidade. O reconhecimento deste fato deve ser levado em conta em qualquer mudança que se pretenda fazer do sistema de avaliação.

- D) No julgamento pelos pares, a informação fundamental é a produção científica. Tal é encontrada no curriculum vitae de cada pesquisador. Quero aqui registrar meu aplauso pela criação e manutenção da Plataforma LATES pelo CNPq. Como todos bem sabem, a plataforma está sendo utilizada por todas as instituições, seja nos julgamentos dos pedidos de auxílio ou de bolsa, seja nas decisões para contratação de docentes, seja na escolha de referees, e a cada dia surgem novas utilizações dela. De fato ela tornou-se um instrumento tão importante que é difícil imaginar como podemos sobreviver anteriormente sem ele.
- E) Na comparação da produção científica de pesquisadores quero enfatizar o fato, conhecido por todos, de que o simples número de artigos publicados tem pouco significado, a não ser para pesquisadores de uma mesma e específica subárea. Cada área do conhecimento possui uma velocidade de publicação própria e, muitas vezes, é preciso subdividir a área em subáreas para se ter a homogeneidade que permita tornar relevante este parâmetro simples que é o número de trabalhos publicados. Por exemplo, na área de matemática se espera, com base na experiência internacional, que um pesquisador da subárea de Geometria publique em média um artigo por ano, enquanto que a expectativa é de um número três vezes maior para a subárea Matemática Aplicada. Algumas universidades, na análise feita para concessão de bolsas de IC, no passado, cometeram o erro de considerar como parâmetro fundamental e único o número de artigos publicados. Tal criou enormes distorções com a supressão de bolsas, por exemplo, para Matemática e Filosofia.
- F) Na busca de julgamentos mais justos a comunidade científica e as instituições têm buscado, ao longo dos anos, formas de tentar medir a qualidade do que esta sendo publicado. O julgamento utilizando referees foi a maneira mais efetiva encontrada, até agora, como resposta a esta busca. A idéia é de que um pesquisador que seja da mesma área do projeto em avaliação, será capaz de destacar, nos currículos dos envolvidos, as contribuições realmente importantes,

será capaz de separar os trabalhos que abrem novos caminhos daqueles que apenas aplicam técnicas conhecidas em situações semelhantes as já estudadas, enfim, será capaz de qualificar as contribuições de vários pesquisadores da sua área. A experiência demonstrou que, para se obter resultados confiáveis nos julgamentos pelos pares, é importante que se utilizem as opiniões de vários pesquisadores em um mesmo julgamento.

- G) A prática do uso de referees revela problemas operacionais bastante complexos. O primeiro deles é da escolha dos referees. Geralmente os que militam nas instituições, cuidando da burocracia envolvida no processo de avaliação, não têm como tomar tal decisão. No passado os pesquisadores envolvidos nos comitês que julgavam os pedidos de bolsas e auxílios entregavam listas de possíveis referees, indicando suas áreas de atuação. Na atualidade acredito que as informações contidas nos currículos disponíveis na Plataforma LATES estejam sendo utilizadas para permitir a criação de tais listas. O segundo problema é da efetividade dos referees. Alguns respondem rápido, outros precisam ser cobrados várias vezes. Alguns escrevem bastante em seus relatórios, outros são lacônicos. Há uma gama de problemas relativos ao funcionamento do sistema de referees que qualquer pessoa envolvida com a operacionalidade do sistema pode relatar. Este também não é um problema nacional sendo alvo de preocupações constantes dos editores de revistas científicas, por exemplo. A solução brasileira para tal problema é a de registrar quais são os referees que atendem com prontidão as solicitações da instituição e que fazem um bom trabalho, descartando-se os demais. Embora esta seja uma solução pragmática, ela leva fatalmente a que se utilizem referees de áreas apenas afins a dos pesquisadores que estão sendo avaliados, destruindo em parte a base sobre a qual o sistema de avaliação pelos pares se assenta.
- H) Uma forma encontrada para avaliar a produção de um pesquisador quando não se dispõe de especialistas na sua subárea é a de verificar a qualidade das revistas onde ele publicou seus artigos. Por exemplo, um artigo publicado no *Annals of Mathematics*, a revista de maior prestígio na comunidade matemática, tem um peso muitas vezes maior do que um artigo publicado em qualquer outra revista. Normalmente a existência de artigo publicado no *Annals* é mencionada em destaque nas avaliações. Por outro lado, existem centenas de revistas publicando artigos de matemática no mundo. O IMPA que tem uma biblioteca do bom nível registra mais de 600 títulos de revistas de matemática. O *Mathematical Reviews* registra cerca de 1200 títulos. Mas, a realidade é que existem mais de 2000 títulos a serem considerados. A questão prática se coloca é: como classificar tais revistas. Tal pode significar trocar um problema mais simples por um muito mais complicado. A CAPES, para subsidiar a avaliação dos cursos de PG, criou o QUALIS que nada mais é do que uma tal classificação em categorias, elaborada com a participação da comunidade científica brasileira. Todas as versões do QUALIS foram duramente criticadas no seio da comunidade. A

última, no caso da Matemática, foi elaborada com uma participação muito ampla e, embora tenha recebido críticas, foi a que menos gerou reações exaltadas. Mas há uma concordância geral de que o QUALIS ainda não pode ser considerado como algo acabado.

- I) A realidade é que existe uma tendência global de que se passe a utilizar critérios “transparentes e objetivos” nos processos de avaliação da qualidade da produção científica. Na prática isto tem conduzido ao uso crescente de parâmetros bibliométricos, sobretudo estatísticas relativas à citações. Esta tendência global se reflete em nosso País, naturalmente, embora sua penetração ainda seja muito variável segundo a área científica. Em particular, os Comitês de Matemática e Estatística da CAPES e do CNPq estão entre os mais prudentes quanto ao uso de critérios numéricos nos seus julgamentos.
- J) Em 2008 foi publicado um documento produzido pela iniciativa de três instituições, entre as quais a União Internacional de Matemática, um relatório de uma comissão que estudou profundamente o assunto e, pelo menos no caso da Matemática, tentou responder a duas perguntas: Até que ponto as estatísticas de citações realmente medem relevância e qualidade? A conclusão do relatório foi a de que não temos garantias de que elas meçam nem a relevância nem a qualidade. De fato, Thomson Scientific, que produz a estatística de citações e de fatores de impacto, diz textualmente:

“Thomson Scientific does not depend on the impact factor alone in assessing the usefulness of a journal, and neither should anyone else. The impact factor should not be used without careful attention to the many phenomena that influence citation rates, as for example the average number of references cited in the average article. The impact factor should be used with informed peer review.”

- K) O fator de impacto vem sendo usado, implícita ou explicitamente, para comparar trabalhos e seus autores. Se o fator de impacto da revista A é maior do que o da revista B, então certamente um trabalho publicado na revista A deve ser superior a um trabalho publicado na revista B, e o autor que publicou em A deve ser superior ao que publicou na revista B. Em vários casos este tipo de argumento se estende para comparar departamentos, grupos de pesquisa e cursos de Pós-Graduação. De fato, um mesmo fascículo de uma revista pode conter artigos de grande profundidade e importância e artigos que representem apenas pesquisa de rotina. Já não é claro que o número de citações represente um indicador preciso da qualidade de um artigo. Quando se trabalha com a média das citações então o número obtido tem muito pouco a ver com qualquer artigo em particular publicado no universo no qual se calcula a média.
- L) O fator de impacto é calculado levando em conta o número de citações nos dois primeiros anos após o ano de publicação. No caso da Matemática o maior

número de citações ocorre depois do terceiro ano e o grau de obsolescência de qualquer artigo é quase zero em todos os casos. Significa dizer que artigos muito antigos podem ser citados em artigos atuais. Por exemplo, o ressurgimento do interesse pelo estudo das superfícies de curvatura média constante levou a um grande número de citações, em artigos da década de 60, de artigos do final do século XIX e início do século XX. De fato, no caso da Matemática a qualidade de uma contribuição científica só pode ser medida a partir de uma retrospectiva histórica.

M) Existem complexas motivações para citação que ainda não são bem entendidas e, por esta razão, as referências feitas nos trabalhos não podem ser simplesmente interpretadas como indicadores de relevância. Entre as várias razões para se incluir uma citação em um artigo, mencionadas na literatura, estão:

- a. Prestar homenagem aos pioneiros
- b. Dar crédito para trabalhos relacionados
- c. Identificar metodologia, equipamento, etc.
- d. Oferecer leitura básica
- e. Ligar o trabalho ao main stream do desenvolvimento da área.
- f. Retificar o próprio trabalho
- g. Retificar o trabalho de outros
- h. Analisar trabalhos anteriores
- i. Sustentar declarações
- j. Informar aos pesquisadores de trabalhos futuros
- k. Fazer propaganda de seus próprios trabalhos
- l. Dar destaque a trabalhos pouco disseminados, inadequadamente indexados ou desconhecidos
- m. Validar dados e categorias de constantes físicas e de fatos, etc.
- n. Identificar idéias originais das quais uma idéia ou um conceito derivam
- o. Contestar trabalhos ou idéias de outros
- p. Debater a primazia das declarações dos outros

N) Por outro lado, o uso de tabelas do tipo QUALIS tem tido um grande impacto sobre os periódicos nacionais. No penúltimo QUALIS a classificação do Boletim da Academia Brasileira de Ciências para Matemática ficou em um nível baixíssimo. A reação imediata da comunidade, particularmente dos jovens, foi a de suspender o envio de novos artigos para tal revista. Ao lado disto, a revista havia passado por uma grande modificação que a havia colocado em um patamar realmente moderno e de alta qualidade. Não preciso dizer que o resultado foi desastroso para os Anais no que concerne a sessão de Matemática. Felizmente

ele publica artigos em várias áreas do conhecimento e pode subsistir a este impacto.

- O) Por estas razões, enquanto se reconhece que "a avaliação precisa ser prática e, por essa razão, estatísticas de citações certamente farão parte do processo", o que se procura é alertar contra o uso excessivo de parâmetros tais como o índice de impacto ou o fator h, "que dão apenas uma visão limitada e incompleta do que é qualidade", insistindo que "a pesquisa é demasiado importante para que o seu valor possa ser medido por meio de uma ferramenta única, e grosseira".
- P) Parece-me que o futuro será guiado principalmente pela busca de um melhor entendimento do que deve ser uma avaliação e pela busca de novos instrumentos, sem descartar o básico que é a avaliação pelos pares. Nesta busca, as instituições, que são as maiores interessadas em ter procedimentos objetivos e claros de avaliação deverão provocar reflexões da comunidade em todos os seus nichos, com a participação das sociedades científicas, no sentido de se obter alterações no imaginário desta comunidade que permita tornar realidade reais mudanças. O trabalho dos comitês assessores, que é de fato um trabalho institucional, deve ser cada vez mais baseado na avaliação pelos pares. Deste ponto de vista a escolha dos pareceristas torna-se então a missão mais importante dos membros dos comitês assessores.

OBRIGADO