

### Mudanças necessárias na avaliação da pós-graduação brasileira

Necessary changes in the evaluation of graduate programs in Brazil

Cambios necesarios en la evaluación del posgrado brasileño

Rita Barradas Barata<sup>(a)</sup>

<rita.barradasbarata@gmail.com> 

<sup>(a)</sup> Faculdade de Ciências Médicas, Santa Casa de São Paulo. Rua Dr. Cesário Motta Jr., 61. São Paulo, SP, Brasil. 01221-020. Ex-Diretora da Avaliação da CAPES.

É inegável a contribuição que a avaliação dos programas de pós-graduação, instituída pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), prestou ao desenvolvimento da formação pós-graduada e da produção científica no país. No entanto, como em qualquer processo avaliativo, após quase vinte anos de aplicação, durante os quais pequenas modificações incrementais e operacionais foram sendo agregadas, o atual modelo precisa ser revisto.

A necessidade de revisão dos pressupostos e dos instrumentos da avaliação decorre da observação, por parte dos coordenadores de área, dos consultores que atuam na avaliação, e dos dirigentes da CAPES, de diversos efeitos indesejados. Entre eles podemos citar: a multiplicação artificial de programas em uma mesma instituição, denotando grande fragmentação dos grupos acadêmicos que enfraquece e compromete a formação e a produção científica na direção da extrema especialização; a dificuldade em ajustar os critérios aos contextos variados de oferta; o foco predominante na produção científica em detrimento

dos aspectos de formação; o uso cada vez mais acríptico dos indicadores quantitativos; a tendência de avaliar o desempenho dos programas por meio do desempenho individual dos docentes ao invés de tomar o programa em si como unidade de análise; o engessamento dos programas que se guiam mais pelo alcance dos critérios do que pela busca permanente da qualidade.

O primeiro aspecto que merece reflexão por parte da comunidade acadêmica é a concepção de avaliação que subjaz aos critérios adotados. Atualmente, a avaliação está baseada na comparação de desempenho entre os programas, gerando uma estratificação ou *ranking* entre os mesmos. Embora não exista uma proporção preestabelecida para a atribuição de notas, como muitos parecem acreditar, basear-se em comparações torna o sistema todo menos previsível, na medida em que a nota recebida por um programa não depende apenas dele mesmo, mas decorre do desempenho do conjunto. Um programa cujo desempenho tenha melhorado ao longo do quadriênio, não necessariamente terá sua nota aumentada, bastando para isso que a melhora observada para os demais tenha sido mais acentuada.

Alternativamente, ao sistema de estratificação, a avaliação poderia estar baseada no alcance de requisitos típicos para cada nota, de modo a tornar a avaliação mais previsível e dependente apenas do desempenho do próprio programa. Se, por um lado, essa alternativa possa parecer mais adequada para a gestão dos programas, ela poderia desestimular a renovação e o esforço de superação, resultando em estagnação do sistema como um todo. Desse modo, a avaliação tenderia a ser sempre mais conservadora, sob pena de estabelecer “tipos” inalcançáveis.

Outro ponto importante na reformulação do processo de avaliação é a necessidade de se equilibrarem abordagens quantitativas e qualitativas. Avaliar pressupõe atribuir valor àquilo que se está analisando. Não há como escapar dessa tarefa atribuindo, aos indicadores quantitativos, uma posição de neutralidade. Dado o tamanho do Sistema Nacional de Pós-graduação (SNPG), é inevitável o recurso a indicadores quantitativos, porém são necessárias, além deles, avaliações qualitativas que permitem diferenciar situações que podem gerar indicadores quantitativos semelhantes a partir de condições muito diversas e, portanto, com significados distintos.

Este ponto se entrelaça à necessidade de se distinguirem tipos e modalidades de oferta que exigem avaliações apropriadas a suas características. Hoje, por exemplo, programas muito grandes, contando com mais de cem docentes, são avaliados sob os mesmos parâmetros que programas muito pequenos, contando apenas com dez a 12 docentes. Certamente, haverá distorções que não serão capazes de refletir corretamente o desempenho de uns e outros.

Entre os documentos elaborados para subsidiar a Comissão Nacional de Acompanhamento do Plano Nacional de Pós-graduação (PNPG) 2011-2020<sup>1</sup> na proposição de uma nova avaliação, o documento encaminhado pela Academia Brasileira de Ciências (ABC) traz várias ideias interessantes, uma delas relacionada mais de perto com essa heterogeneidade existente no SNPG. Os acadêmicos propuseram que os programas pudessem autodefinir sua vocação, possibilitando, assim, a criação de “clusters” de programas assemelhados, aos quais se aplicariam critérios de avaliação mais ajustados. Os grupos sugeridos foram: pesquisa básica, pesquisa básica em temas estratégicos, pesquisa aplicada na área social e pesquisa aplicada na área tecnológica.

O documento final da Comissão Nacional, recentemente aprovado pelo Conselho Superior da CAPES (10/10/2018)<sup>2</sup>, propõe adotar um modelo de avaliação multidimensional para contemplar diferentes aspectos, sem que fosse necessário resumir-los em uma única nota. A abordagem multidimensional permitiria acomodar a diversidade de vocações e objetivos dos programas. As dimensões sugeridas, a título de exemplo, foram: formação de recursos humanos, internacionalização, produção científica, inovação e transferência de conhecimento, impacto e relevância econômica e social.

Há, pelo menos, três aspectos cruciais que atualmente estão ausentes ou subvalorizados na avaliação e que deveriam fazer parte da nova proposta, todos eles abordados pela Comissão Nacional: a questão da formação, a autoavaliação e o impacto econômico e social.

A avaliação atual está profundamente calcada na produção intelectual do programa que corresponde a praticamente 70% da nota (30% a 40% da nota atribuída à produção intelectual dos docentes e 30% a 40% atribuída à produção dos discentes). As questões relacionadas à formação estão subavaliadas, uma vez que a estrutura de oferta do programa (proposta do programa) não contribui diretamente para a nota, muito embora possam constituir uma trava para o alcance de determinadas notas.

A formação vem sendo avaliada indiretamente pela publicação dos resultados dos trabalhos finais sob a forma de artigos ou livros e da produção imediata (três a cinco anos) de egressos dos programas. Como correlato dessa “desvalorização”, ao longo do tempo, os programas foram diminuindo a carga horária das disciplinas e de outras atividades formativas, resultando na titulação de novos pesquisadores nem sempre com a autonomia e a liderança intelectual desejadas. A ideia de que os pesquisadores podem ser formados quase que exclusivamente nas atividades dos próprios grupos de pesquisa necessita ser revista. Nos programas atuais, praticamente não há espaço para a formação teórica e metodológica ampla, ficando os alunos limitados aos aspectos mais técnicos da reprodução de pesquisas. A formação para docência também é bastante relegada, ainda que os estudos com egressos mostrem que a maioria deles irá se inserir em atividades docentes no ensino superior.

Um ponto que vem sendo debatido internacionalmente por diversas academias de ciência é a necessidade de se agregarem componentes de humanidades e arte à formação científica, recolocando a questão da formação de intelectuais para além da preparação técnica de cientistas. Articulada a essa necessidade de formação mais ampla, coloca-se a questão da diversidade de perfis dos docentes, hoje, praticamente impedida pelos critérios de avaliação.

A nova avaliação precisará considerar a formação dos sujeitos (docentes, pesquisadores ou profissionais) em sua plenitude, não se restringindo apenas aos produtos científicos ou tecnológicos. Nesse sentido, avaliar a trajetória dos egressos ganha relevância central para bem avaliar os programas, ainda que pouco se saiba como fazê-la.

A autoavaliação sistemática pelo próprio programa deve também passar a ocupar posição de maior importância no processo avaliativo final. É importante que docentes e discentes elaborem uma autoavaliação que possa ser compartilhada com os consultores da comissão de avaliação e permita contextualizar o programa e sua trajetória. Caberá ao programa indicar, para a comissão avaliadora, suas virtudes e

eventuais fragilidades. Quem foram seus melhores alunos no período? Quais as teses e dissertações que merecem destaque? Qual foi sua produção intelectual mais relevante? Que contribuições fundamentais o programa teve, em sua região, para o país ou internacionalmente?

A indicação da melhor produção do programa pode permitir, às comissões, avaliarem qualitativamente os produtos e egressos ao invés de considerar a produção total, cujo volume tem tornado impossível, às comissões, aquilatarem a real contribuição do SNPG para o desenvolvimento do país e a posição relativa dos programas para além dos indicadores quantitativos.

Do mesmo modo, a avaliação do impacto e da relevância social e econômica dos programas está atualmente subavaliada no quesito de inserção social. Os atuais instrumentos de registro não permitem às comissões realmente qualificarem as contribuições que o conhecimento produzido e os egressos formados podem ter tido no âmbito local, regional, nacional ou internacional.

Todas essas mudanças, que são necessárias e precisam ser viabilizadas, dependem de modificações em aspectos operacionais e nos instrumentos utilizados para a avaliação.

Dependendo dos consensos construídos com a comunidade acadêmica em torno dos aspectos conceituais, será necessário redesenhar instrumentos e redefinir procedimentos.

Um aspecto que necessita de mudança diz respeito à escala de notas hoje utilizada. Muito embora as notas variem entre 1 e 7, na prática, para a maioria dos programas em atividade, a escala se resume a apenas três pontos (notas 3, 4 ou 5), uma vez que as notas 1 e 2 significam o descredenciamento do programa, e as notas 6 e 7 denotam excelência. Uma alternativa seria adotar as notas de 1 a 5 como notas válidas, e não atribuir notas tanto para o descredenciamento quanto para a classificação da excelência. Maior amplitude da escala daria, aos avaliadores, condições mais adequadas de refletirem os avanços dos programas.

Outro ponto, praticamente consensual em todos os fóruns de discussão da avaliação, diz respeito à inadequação da Ficha de Avaliação. Com seus itens e pesos muito amarrados, ela representa um entrave a uma avaliação mais qualitativa.

O Coleta, principal instrumento de registro das informações pelos programas para uso pelas comissões, precisará ser ajustado. Vários módulos precisarão ser substituídos ou reformulados, tornando-os mais úteis. Precisam de reformulação mais urgente: o módulo proposta, que deverá ser ajustado para diferentes modalidades (acadêmica ou profissional) e tipos de oferta (programa isolado, em associação ou em rede); o módulo de livros, que deverá conter informação básica para análise global da produção, e o módulo de produção técnica para a avaliação dos programas profissionais.

Os estudos de egressos precisam ser mantidos e aprimorados para possibilitarem melhor avaliação do processo de formação. Atualmente, a CAPES dispõe de pesquisas realizadas pelo Centro de Gestão e Estudos Estratégicos (CGEE) a partir de *linkage* da base de dados dos egressos e base de dados da Relação Anual de Informações Sociais (RAIS) do Ministério do Trabalho; mas, para algumas áreas de avaliação, a cobertura é relativamente menor, uma vez que só estão incluídos os indivíduos com trabalho formal no país. Além disso, as informações só permitem avaliar: a empregabilidade, a geração de renda e a mobilidade espacial e entre áreas do conhecimento entre os cursos de mestrado e doutorado.

Finalmente, será necessário redefinir o Qualis periódicos, no sentido de torná-lo uma ferramenta mais homogênea entre as áreas e capaz de sinalizar as características gerais de produção dos programas. Hoje, os critérios adotados pelas diferentes áreas são incomensuráveis, dificultando o diálogo entre elas e produzindo muitos ruídos entre: editores, autores, agências de fomento, políticas institucionais nas instituições de ensino superior, entre outras.

Se os periódicos nos quais a produção está sendo divulgada, estão sendo usados como substitutos para um processo avaliativo de *peer review* pelas próprias comissões, não há como justificar como o mesmo periódico receba classificações tão díspares entre as áreas.

Uma das possibilidades de ordenamento, sem ignorar as particularidades de cada área de conhecimento e entre as áreas de avaliação, seria construir um Qualis em que cada periódico seria atribuído a uma única área de avaliação de acordo com seu escopo de publicação, e cada área de avaliação classificaria apenas os periódicos em sua área de abrangência. O segundo passo seria estabelecer indicadores bibliométricos comuns, como o índice h, fator de impacto ou fator de impacto normalizado (SNIP). O ideal é combinar métricas que avaliem aspectos distintos do impacto. Os pontos de corte (percentis) para os indicadores deveriam ser estabelecidos a partir de uma listagem exaustiva das publicações em cada área, e não a partir da produção dos programas para posicionar corretamente a produção nacional vis a vis a produção mundial. Para as publicações das ciências humanas e sociais, o Google Scholar pode ser um substituto a outras bases indexadoras nas quais essa produção está sub-representada.

Resumidamente, uma nova avaliação precisa estar mais focada nos aspectos de formação, e não apenas na produção de conhecimentos; combinar indicadores quantitativos e qualitativos; possibilitar melhor contextualização dos programas; incluir a autoavaliação; focalizar a melhor produção ao invés da produção global, seja de egressos seja de conhecimentos; valorizar a relevância social e econômica dos programas, e evitar análise individualizada por docentes, priorizando o programa como um todo.

Este é o desafio que a comunidade acadêmica deve enfrentar nos próximos anos para reformular o sistema de avaliação e continuar auxiliando o desenvolvimento do país.

### Direitos autorais

Este artigo está licenciado sob a Licença Internacional Creative Commons 4.0, tipo BY ([https://creativecommons.org/licenses/by/4.0/deed.pt\\_BR](https://creativecommons.org/licenses/by/4.0/deed.pt_BR)).



### Referências

1. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Plano Nacional de Pós-Graduação - PNPG 211-2020 [Internet]. 2018 [citado 10 Nov 2018]. Disponível em: <http://www.capes.gov.br/plano-nacional-de-pos-graduacao>



2. Brasil. Ministério da Educação. Proposta de Aprimoramento do Modelo de Avaliação da PG. Documento Final da Comissão Nacional de Acompanhamento do PNPB 2011-2020 [Internet]. Brasília, DF: MEC; 2018 [citado 10 Nov 2018]. Disponível em: [https://www.capes.gov.br/images/stories/download/conselho-superior/18102018\\_PNPB\\_CS\\_Avaliacao\\_Final\\_CS\\_FINAL\\_17\\_55.pdf](https://www.capes.gov.br/images/stories/download/conselho-superior/18102018_PNPB_CS_Avaliacao_Final_CS_FINAL_17_55.pdf)

Submetido em 12/11/2018.

Aprovado em 12/11/2018.