

Desafios institucionais e perspectivas para avaliação responsável de pesquisa na educação superior brasileira¹

O que é avaliação?

A avaliação pode ser aplicada com diferentes propósitos. Pode servir para certificar as instâncias de governança sobre o desempenho em atividades específicas em uma instituição. Também pode ser utilizada para aferir da concepção ao resultado de um plano; para levantar a contribuição de um indivíduo específico na consecução de um objetivo institucional; para determinar se um membro da comunidade de pesquisa deve ser promovido ou premiado; ou ainda se algum aspecto de planejamento carece de revisão ou necessita uma correção de rumo. É possível utilizar sistemas avaliativos para determinar se um projeto de pesquisa deve receber financiamento, ou ainda para aferir as contribuições dos resultados de um projeto de pesquisa para a sociedade de forma geral.

O que é avaliação responsável?

Responsabilidade em avaliação implica na adoção de práticas orientadas por um conjunto de princípios, que permitam relativizar as medidas quantitativas simplificadoras para avaliar a qualidade da pesquisa. Uma avaliação responsável busca oferecer aos pesquisadores uma ampla variedade de indicadores para os impactos econômicos, sociais, culturais, ambientais ou em termos de formulação de

¹ Este documento é um resumo dos principais argumentos oriundos da interação entre o projeto Métricas e a iniciativa de engajamento comunitário da Declaração de São Francisco (DORA), apoiada pela American Society of Cell Biology. O projeto que deu origem a este documento reuniu uma comunidade de especialistas em avaliação numa série de atividades com o objetivo de refletir, identificar a situação presente, as perspectivas e oportunidades para o futuro. O primeiro *workshop* foi realizado em julho de 2022. Na ocasião, os participantes foram convidados a comentar suas experiências pessoais, pesquisas e o resultado de uma enquete realizada entre signatários da DORA, vinculados a instituições brasileiras. Às respostas do questionário foram aplicadas as rubricas Space, para diagnosticar as tendências de opinião da comunidade sobre avaliação de pesquisa nas instituições brasileiras. Ao fim desse processo, foram identificadas três áreas de ação prioritárias sobre avaliação responsável: Conscientização das comunidades sobre avaliação responsável; Treinamento e capacitação para avaliação responsável; Execução e aprimoramento de processos de avaliação. Esses três temas foram trazidos ao debate durante uma atividade pública realizada no dia 19 de agosto de 2022, com registro em vídeo (disponível em https://www.youtube.com/watch?v=auu81laD_mY). O objetivo principal desse segundo evento público foi identificar perspectivas, oportunidades e prioridades para a superação dos desafios apontados no primeiro evento.

políticas resultantes de seus projetos de pesquisa. Uma avaliação responsável deve considerar uma mistura de métodos de avaliação qualitativos e quantitativos. Essa combinação deve ser calibrada para descrever tanto a qualidade quanto o impacto de maneira mais completa possível, de forma a evitar as tendências inerentes aos métodos de avaliação empregados e, ao mesmo tempo, que permita a comparação entre diferentes projetos e programas de pesquisa. A melhor forma de criar consciência sobre práticas de avaliação responsável é debater dentro de cada comunidade específica as ferramentas e documentos, tais quais a Declaração de São Francisco ou o Manifesto Leiden.

Conscientização das comunidades sobre avaliação responsável

As questões fundamentais para elevar a consciência sobre avaliação responsável nas instituições são: Por que, o que, como e quando avaliar?

Por que as universidades devem avaliar de forma responsável?

Para que um processo de avaliação seja efetivo, as universidades devem ter clareza das razões e dos objetivos do processo avaliativo em questão. Avaliações promovem certos valores e comportamentos projetados para o futuro, além de fornecer respostas no presente, ao mesmo tempo que permitem analisar o desempenho em retrospecto. Os objetivos dos sistemas de avaliação podem ser de variadas naturezas, tais como: identificar e apoiar pesquisadores destacados com potencial de auferir reconhecimento internacional em suas áreas; fomentar a inovação e identificar potenciais empreendimentos com capacidade de fundar novos negócios baseados em conhecimento; ou concentrar-se na qualidade da formação de seus egressos, que farão a diferença na sociedade, contribuindo com a excelência em suas ocupações profissionais. *Cada um desses objetivos exige desenhos específicos de sistemas de avaliação que devem ser observados pelas universidades em sua concepção.*

Considerando-se a diversidade de áreas do conhecimento, tipos de instituições, trajetórias de carreira e fatores socioeconômicos observados nos sistemas de educação superior brasileiros, *uma ampla variedade de modelos de avaliação deve ser considerada e estabelecida* para assegurar que a diversidade e a heterogeneidade de missões, valores e resultados esperados sejam respeitados.

Avaliação da qualidade da produção científica é um fenômeno complexo e em constante mutação. Devemos evitar, portanto, confiar em critérios fixos para realizar esse julgamento. Além disso, como nos explica a lei de Goodhart, sempre que uma

métrica se torna um objetivo em si, ela deixa de ser uma boa métrica². A avaliação responsável não deve ser entendida como a busca pelo conjunto perfeito de indicadores, senão como um ciclo constante de validação e revisão de processos. É necessário que as universidades desenvolvam mecanismos de verificação constante do desempenho e da qualidade de cada um de seus processos e metodologias de avaliação.

Uma avaliação responsável deve oferecer *feedback* a fim de incentivar a assunção de riscos, aceitar a falha como parte do processo da descoberta, ao reconhecer a incerteza e a ignorância como componentes importantes do processo de criação de um conhecimento significativo. Dessa forma, uma abordagem enfática de ética e integridade, especialmente mediada pelos códigos de ética das universidades, é um componente importante nesse processo de aumento da consciência das comunidades acadêmicas a respeito da avaliação responsável.

Quais ações e medidas práticas devem ser tomadas?

Para implantar qualquer mudança, toda comunidade acadêmica deve estar informada e consciente a respeito dos fundamentos e implicações de uma avaliação responsável, bem como das recomendações inerentes. Ao ingressar em um sistema de produção de ciência, um pesquisador passa a submeter-se à avaliação do seu desempenho como indivíduo, parte de um grupo de pesquisa, como membro de um departamento, de uma escola ou de uma universidade, concordando ou não com os critérios dessa avaliação. O conhecimento prévio sobre DORA, Leiden ou outras iniciativas similares ajuda a comunidade a familiarizar-se com os argumentos e vocabulário que viabilizam uma discussão mais sofisticada em torno de práticas de avaliação.

Algumas estratégias de difusão e conscientização das comunidades podem preceder, ou ainda seguirem imediatamente a adesão à DORA por parte das instituições. Estudantes podem ser estimulados a perceber a importância de processos de avaliação de cursos. Pesquisadores em início de carreira devem ser informados sobre os princípios de avaliação responsável, tal qual os ingressantes na instituição e também os pesquisadores mais experientes.

Além da ausência de conhecimento a respeito da DORA ou de documentos similares, é patente a falta de experiência prévia e capacitação de avaliadores e avaliados. O caminho para a avaliação mais responsável demanda programas de treinamento e educação adicional para garantir a consolidação de uma cultura de produção

² "When a measure becomes a target, it ceases to be a good measure" (M. Strathern, "Improving Ratings: Audit in the British University System", *Eur. Rev.*, vol. 5, pp. 305–321, 1997).

científica orientada para o impacto, e de um ambiente propício para o estabelecimento de uma avaliação responsável.

Como as universidades podem implantar processos de avaliação mais responsáveis?

Princípios e orientações inerentes aos processos de avaliação

Princípios de avaliação responsável precisam permear todos os processos sobre os quais incidem práticas avaliativas. Procedimentos claramente comunicados a partir da administração central e que explicam como esses princípios se aplicam às chamadas públicas para contratação de docentes e pesquisadores, aos processos de promoção, avaliação departamental e avaliação para fins de acesso a recursos financeiros.

Uma vez estabelecido o compromisso com a avaliação responsável, a instituição deve conectar essas novas ideias com valores já consolidados em suas atividades. O principal, para que a ciência avance, expressa-se no desejo de transcender as barreiras do conhecimento, contribuindo para a melhora da qualidade de vida na sociedade por meio de sua disseminação e aplicação. A avaliação responsável incorpora em si esses valores.

Os objetivos e ambições da avaliação devem estar claros para assegurar que os pesquisadores entendam, acreditem e confiem nos seus propósitos, e que sempre há abertura para comentários e observações da comunidade visando à melhoria contínua dos processos de avaliação. Estes objetivos devem estar vinculados a objetivos institucionais específicos e claros.

O objetivo desejado desse esforço é prover uma *avaliação que faça sentido* para as comunidades e que respeite os valores intrínsecos à atividade científica. Alcançar esse tipo de consciência exigirá o *envolvimento de vários atores institucionais*. Enquanto a liderança da administração das universidades é necessária para atuar em nível institucional junto às suas comunidades docentes e discentes, é fundamental que a avaliação responsável também seja discutida e defendida por *sociedades e academias científicas, agências de financiamento e outras organizações científicas*.

O uso de indicadores e métricas é necessário para estabelecer bases de comparação entre unidades e para aferir seu desempenho ao longo do tempo. O uso de todas esses parâmetros, no entanto, traz consigo limitações e deficiências. Qualquer métrica ou indicador escolhido para representar uma dimensão de

desempenho implicará na possibilidade de distorção de comportamentos. Quanto maior for a dependência em relação a um conjunto limitado de indicadores, maior será essa tendência de distorção. O desafio de construir as capacidades para lidar com esse problema passa por *discutir e aumentar a consciência dessa tendência dentro das comunidades científicas*, dando condições aos avaliadores de compreender e interpretar o desempenho de acordo com os entendimentos da comunidade.

Quais incentivos institucionais podem ser oferecidos para encorajar boas práticas de avaliação?

Há um equilíbrio que deve ser constantemente calibrado e perseguido. De um lado, interesses individuais e, do outro, as prioridades da instituição. Quando indicadores escolhidos estão em alinhamento com as expectativas pessoais e institucionais, é esperada uma sinergia entre esses critérios.

Práticas empregadas em instituições brasileiras que podem servir de modelo

Universidades precisam conciliar objetivos pessoais, locais e institucionais em seus modelos de avaliação. O modelo de súmula curricular adotado pela Fapesp solicita ao pesquisador priorizar as cinco realizações mais importantes da carreira. Esse exercício permite, em nível mais amplo, *privilegiar os aspectos de excelência, qualidade das carreiras individuais, distanciando-se de uma perspectiva de saturação e hiperprodutividade* e pode ser adotado como formato desejável para avaliação individual.

Memoriais narrativos e descritivos já são instrumentos de uso comum em concursos de livre-docência e outros meios de avanço nas carreiras dos docentes. Esses documentos mais longos e detalhados proporcionam mais liberdade ao pesquisador para argumentar sobre suas contribuições à ciência e podem oferecer a base para futuras metodologias de avaliação. Há de se considerar, no entanto, que esses memoriais são pouco explorados como ferramenta de avaliação. Em alguns casos, a preparação de um memorial carece de normas estabelecidas pelas instituições para sua confecção, tanto para seu formato quanto para o conteúdo, bem como com respeito à sua extensão.

Acompanhamento das atividades de egressos como medida de impacto do ensino é muito mais relevante do que qualquer outra metodologia comumente empregada para avaliação de atividades de ensino. No entanto, há de se buscar o equilíbrio pelo entendimento de que a simples soma dos impactos individuais deve ser ponderada por fatores intermediários, como por exemplo características institucionais particulares ou de condições socioeconômicas do entorno.

Qual o momento adequado para uma universidade planejar e executar um plano de adoção de avaliação responsável?

Mudanças devem ser planejadas e implantadas de forma incremental, entre ciclos de avaliação, afastando-se progressivamente da dependência de indicadores bibliométricos em direção a métodos mistos de aferição. O êxito em cada movimento incremental dependerá de amplo apoio da comunidade acadêmica, e no processo, avaliadores e avaliados devem ter satisfeita a percepção de que seus esforços tenham sido válidos e reconhecidos.

Este tipo de mudança requer planejamento de longo prazo, *com muita clareza na definição de objetivos alinhados com as ambições da instituição*, além da consciência de que dificilmente será possível alcançá-los em um único ciclo de mudança. Alterações drásticas em processos de avaliação raramente são bem-sucedidas, além de contribuírem para degradação do ambiente acadêmico e descontentamento da comunidade, caso esta relação entre as mudanças e as visões de longo prazo da instituição sobre como ela pretende avaliar a qualidade científica não estiverem claramente estabelecidas.

Repensar os ciclos temporais da avaliação é importante. A avaliação responsável demanda mais tempo do que aquelas baseadas meramente em dados quantitativos. Essa afirmação é válida tanto para quem tem a responsabilidade de submeter evidências quanto para os que avaliam. Ciclos curtos de avaliação encorajam o uso de métodos antiquados. Metodologias mistas, no entanto, demandam mais tempo e, conseqüentemente, reflexão para que os pesquisadores coletem evidências do impacto de suas atividades de pesquisa.

Novas formas de avaliação devem evoluir harmoniosamente, evitando uma carga excessiva de trabalho aos pesquisadores. A introdução de novos formatos de relatórios de impacto ou súmulas curriculares deve levar esse aspecto em consideração.

Treinamento e capacitação para avaliação responsável

Para surtir o efeito desejado, *atividades de treinamento e capacitação devem preceder quaisquer mudanças nos critérios de avaliação.* A transparência nos processos de reforma e oferta antecipada de oportunidades de treinamento à comunidade podem diminuir a ansiedade e a resistência a mudanças. Essa postura aumenta a chance de implantar uma mudança cultural de forma efetiva e duradoura.

Identificar os objetos de avaliação deve ser prioridade. *Identificar com exatidão os objetos de avaliação permite evidenciar o melhor formato, a estrutura e os objetivos do processo.* Ter clara consciência a respeito desse objeto, seja ele um curso de graduação, um projeto de pesquisa ou um pesquisador em início de carreira, torna mais fácil à comunidade tomar decisões ou corrigir rumos do sistema de avaliação no futuro.

Três ações concretas para considerar na construção de capacidades em avaliação

- Realizar *workshops* sobre os usos e abusos de indicadores de pesquisa, para elevar o conhecimento específico sobre o tema, visando capacitar corpos de avaliadores para compreender, interpretar e realizar julgamentos qualitativos.
- Indicação de modelos aplicáveis de avaliação para aferição de objetivos específicos nas instituições, alinhados com prioridades institucionais.
- Comunicação franca, promovendo constante debate e revisão de indicadores e processos, visando alinhar as expectativas da comunidade com a realidade institucional.

Treinamento e capacitação de avaliadores

Avaliadores precisam receber treinamento para realizar seu julgamento com base em informações de naturezas qualitativas e quantitativas. Assim, avaliar de forma embasada em variados tipos de evidências de maneira consistente e estruturada.

Treinamento e capacitação de avaliados

O objetivo deste treinamento é oferecer elementos para refletir, escrever e levantar evidências a respeito do impacto da atuação do sujeito na avaliação. Pesquisadores tendem a confiar em indicadores quantitativos ou na elaboração de relatórios pouco substanciados. Para inverter essa tendência, é necessário um esforço de treinamento e capacitação a partir dos primeiros passos da carreira acadêmica, para *fomentar uma cultura de avaliação* que discuta desde a formulação de um plano de pesquisa e sua execução e permita entender os relatórios como o coroamento de um esforço de planejamento de avaliação.

Sistemas avaliativos devem considerar diferentes *níveis de avaliação para selecionar apropriadamente os instrumentos de mensuração*. Avaliações individuais de desempenho de um pesquisador, seja para contratação ou progressão na carreira, demandam diferentes conjuntos de métodos e indicadores do que as avaliações coletivas, de departamentos ou unidades que, por sua vez, diferem do conjunto de metodologias necessárias para aferir o desempenho de toda uma instituição universitária. Enquanto cada nível tem suas peculiaridades e

especificidades que devem ser consideradas em um processo de avaliação responsável, é também importante observar como se dão as interações entre esses diferentes níveis. O desafio é assegurar-se que os resultados aferidos em um nível estejam convergentes com os objetivos definidos nos demais. É possível afirmar que *avaliação é uma atividade holística, que busca o equilíbrio entre anseios individuais e as aspirações institucionais.*

Corpos avaliadores devem carregar a responsabilidade de resguardar a memória institucional e a experiência acumulada em ciclos de avaliação anteriores. À medida que esses grupos são responsáveis pelo ciclo avaliativo presente, devem também engajar-se no planejamento e aprimoramento reflexivo dos ciclos futuros de avaliação. Cabe a esse conjunto de avaliadores resguardar a qualidade desse processo de revisão constante da avaliação em consonância com os objetivos de suas respectivas unidades e preocupar-se com a comparação de seus processos de avaliação com outros semelhantes, em diferentes áreas de conhecimento e diferentes instituições.

É necessário que a avaliação tenha significado para a comunidade. Isto pode ser obtido pela *celebração e valorização de resultados excepcionais*, ou pela indicação precisa de pontos carentes de aprimoramento. Tanto em um caso como no outro, é necessário deixar muito evidentes os pontos levantados pela avaliação para que resulte em *recomendações claras para incorporação em futuros ciclos avaliativos.*

Execução e aprimoramento de processos de avaliação

Cada processo da educação superior ocorre em distintos ciclos temporais. A duração de um curso de graduação, o tempo de vigência de uma posição de pós-doutorado, a duração de um ciclo de avaliação departamental, um projeto de pesquisa, o mandato de uma administração reitoral, todos esses ciclos acontecem de forma assíncrona e, o que é importante considerar, seus impactos são revelados em ciclos também muito diferentes entre si. O impacto de um curso de graduação pode revelar-se ao longo de toda uma vida, enquanto o impacto de um resultado de pesquisa pode aparecer em alguns meses após a sua publicação, ou mesmo demorar anos ou décadas para se revelar. Diferentes áreas do conhecimento frequentemente apresentam características temporais muito diferentes. Enquanto nas humanidades alguns assuntos podem levar décadas de esforços para se consolidarem e outras tantas para ter sua influência difundida, em outros campos emergentes relacionados à tecnologia, como por exemplo na computação, uma pesquisa pode alcançar impacto considerável em questão de semanas.

Para que os objetivos da avaliação sejam alcançados, *é fundamental levar em conta a correta definição dos horizontes temporais*, tanto no planejamento quanto na execução da avaliação.

Um planejamento cuidadoso dos ciclos de avaliação deve *prevenir a repetição de exercícios de avaliação, bem como a sobreposição desnecessária de procedimentos*. A exaustão avaliativa é um fenômeno bem conhecido na educação superior. Ela se expressa na tarefa imposta aos avaliados de submeter de forma repetida e recorrente os mesmos conjuntos de informações para diferentes instâncias e propósitos de avaliação. Minimizar esse retrabalho ajuda a aumentar a aceitação das comunidades aos processos de avaliação.

A evolução da avaliação exige planejamento cuidadoso nos diferentes horizontes de curto, médio e longo prazo. Mudanças repentinas e dramáticas são difíceis, senão impossíveis de serem executadas nas universidades. Assim, objetivos estratégicos de longo prazo determinados por uma visão clara dos alvos da instituição, reforçados com ações de curto e médio prazo, são a forma recomendada de proceder com relação aos processos de avaliação.

Do grupo que observa o processo de avaliação exige-se uma *compreensão dos diferentes horizontes e ciclos temporais*, tendo sempre em mente quais são os aspectos mais apropriados de atenção em cada período, *sempre oferecendo recomendações para ajustes nos ciclos futuros*.

É de vital importância garantir que cada ciclo de avaliação seja devidamente concluído, que não se abandone nenhuma etapa e que mudanças não aconteçam durante a execução de um ciclo de avaliação. Os ciclos podem, no entanto, ser alongados ou encurtados conforme recomendações provenientes de diálogo entre a comunidade.

Os objetivos da avaliação devem ser constantemente discutidos e validados entre cada ciclo. Em razão do entendimento de que objetivos institucionais podem mudar ao longo do tempo, conforme influência de fatores internos ou externos às instituições, *as práticas de avaliação também devem ser constantemente debatidas e repactuadas para refletir as mudanças de prioridades institucionais*. Toda e qualquer revisão desse tipo deve acontecer entre ciclos avaliativos, preparando-a para que entre em vigor sempre no próximo ciclo.

Conclusões

A implantação de processos de avaliação responsável nas universidades brasileiras traz grandes benefícios, como o de melhorar a relação das instituições de ensino

superior, sobretudo as públicas, com o conjunto da sociedade. Processos de avaliação responsável aumentam a capacidade de diálogo e transparência entre as universidades e a sociedade por incorporar valores como a flexibilidade e a representatividade da avaliação das atividades científicas. O alinhamento de valores da ciência, das instituições e dos indivíduos contribui para motivar as comunidades acadêmicas e assegurar maior clareza de valores e princípios expressos nos caminhos trilhados pelas instituições.

Para atingir essa finalidade a adesão a declarações de princípios, a exemplo de DORA e Leiden, assim como a conscientização das comunidades envolvidas em processos de produção de conhecimento, é um primeiro passo importante. O passo imediatamente seguinte é o de garantir processos de treinamento e capacitação à comunidade científica. Trata-se de oferecer mais clareza e capacidade aos produtores de ciência de explicar as intenções, os resultados e os impactos de seus esforços de pesquisa e de assegurar aos avaliadores condições de emitir julgamentos solidamente balizados sobre as atividades dos seus pares.

O planejamento de mudanças em sistemas de avaliação deve ser cuidadosamente estudado em relação às dimensões temporais. Cabe considerar os objetivos dos três horizontes, de curto, médio e longo prazo, tomando cuidado para que toda ação seja implantada em um ritmo adequado. Ciclos de avaliação devem ser planejados evitando sobreposição de atividades e retrabalho. Corpos de avaliadores devem, pelo menos em parte, ser mantidos entre ciclos, evitando renovação total desses grupos, para garantir a manutenção da memória institucional entre ciclos sucessivos, evitando assim que as instituições tenham que iniciar seus sistemas a partir do zero em cada reinício.

Avaliação responsável não é a representação ideal de um único conjunto de indicadores e métricas perfeitos, mas sim um processo contínuo de aprimoramento gradual e revisões reiteradas dos processos de avaliação. Por essa razão, o debate constante e a revisão de procedimentos são requisitos para garantir que a avaliação possa evoluir, sempre refletindo as mudanças nas circunstâncias e nas demandas da sociedade, à qual as nossas instituições servem e que está em constante transformação.