

“TEM QUE MUDAR A FORMA DE AVALIAR”, DIZ EX-DIRETORA DE AVALIAÇÃO DA CAPES

Daniel Garcia



- **“Não tem que avaliar cada pessoa**, um erro enorme da pós-graduação é que não está avaliando programa, está avaliando soma de professores. E cada professor tem que ser igualzinho ao outro, no sistema atual, porque você tem as regrinhas: ‘Eu preciso ter três a cinco alunos’, ‘Eu preciso publicar xis’. Fica todo mundo igualzinho. **Isso é péssimo para a produção de conhecimento**. Para mim, hoje, o problema maior da avaliação está aí”
- “Comecei a perceber que esta avaliação estava muito parametrizada, muito estabelecida em cima de pontos de corte para todas as variáveis. Não é só para a produção, as outras coisas também: quantos alunos por orientador, qual tempo de titulação, **tudo isso estava muito transformado em indicador aritmético**, esquecendo que o mesmo número representa situações completamente diferentes”
- “Tem que mudar a forma de avaliar [os artigos]. Então talvez não possa avaliar tudo. Talvez tenha que levar em conta o que é mais relevante. Por isso a questão da autoavaliação do programa, que precisaria ter um peso maior”. “Substituir a ficha por um roteiro de avaliação, mais livre, onde eu tivesse espaço e não ficasse amarrada nos ‘pesinhos’ e nas ponderações”. “Você avançaria, e **desmontaria essa ideia de que todo mundo precisa fazer 10 artigos, 15 artigos, 20 artigos**. Não. Você precisa fazer um, muito bom”

Pode ser surpreendente para alguns. A autora das frases ao lado, que “tombam” frontalmente com a mentalidade produtivista em voga, não é estranha às searas do sistema de avaliação acadêmica, antes pelo contrário. Trata-se da professora Rita de Cássia Barradas Barata, da Faculdade de Medicina da Santa Casa de São Paulo, que foi diretora de Avaliação da Coordenação de Aperfeiçoamento do Pessoal de Nível Superior (Capes) entre agosto de 2016 e abril de 2018, tendo antes ocupado outros cargos na casa. Rita coordenou a área de Ciências da Saúde do Conselho Técnico-Científico (CTC) da Capes. Também representou o CTC no Conselho Superior da agência, que, como se sabe, é o órgão incumbido de regulamentar a pós-graduação nacional e avaliar sua qualidade, por meio de um sistema de avaliação trienal que atribui notas aos diferentes programas de mestrado e doutorado.

Nesta entrevista à **Revista Adusp**, concedida em junho de 2019, Rita retoma questões que já vinha trabalhando publicamente ao menos desde 2017, quando, numa concorrida reunião com as coordenações de programas de pós-graduação da USP, expôs suas críticas e anunciou medidas iniciais de reforma do sistema — as quais, no entanto, só conseguiria implantar em parte. Em agosto de 2018, em reveladora entrevista à jornalista Camille Cardoso, publicada no blogue **Direto da Ciência**, a professora recorreu à metáfora da **máquina doida**, que ela voltou a citar para esta revista: “Quando todos publicam o dobro pensando na avaliação seguinte, os pontos de corte dobram. Então é ilógico. Se a roda não for desmontada, ela não vai se desmontar sozinha, vai continuar rodando que nem doida. O sentido original da coisa se perdeu. É preciso voltar a olhar a formação, porque os doutores estão saindo mais como técnicos do que como pensadores.”

Que fique claro: a professora está entre os defensores da avaliação (“como é que eu sei que devo pôr mais dinheiro aqui, menos ali?”), mas acredita ser possível um sistema inteligente, antípoda do produtivismo acadêmico, baseado na qualidade, ainda que reconheça ser inexecutável uma avaliação inteiramente qualitativa. Neste modelo alternativo ganhariam espaço a autoavaliação dos programas, mediante a substituição da ficha de avaliação por um roteiro sem engessamento, e a diversidade de perfis.

O problema maior, atualmente, diz Rita, é a uniformização estéril provocada pelo sistema em vigor. “Se tivesse que indicar qual é ‘o’ problema, é esse. Eu tenho que avaliar o programa, o programa tem que ter diversidade, eu preciso ter diferentes perfis. Eu tenho que ter aquele cara que é um ótimo pesquisador, mas é um péssimo docente. Que a gente sabe que existe. Então eu vou pôr os alunos para aprender com ele? E esse que gosta de ensinar vai produzir menos que aquele que faz só pesquisa. Mas os dois são úteis para o programa”. **(Pedro Estevam da Rocha Pomar e Daniel Garcia)**

Revista Adusp. Em que momento a Sra. começou a se dar conta dessas distorções, desses problemas?

RBB. Já no CTC. Porque eu coordenei a área de saúde coletiva por dois períodos. Quando você começa a lidar com o conjunto dos cursos, primeiro da sua área, e entender como a dinâmica ocorre nos diferentes cursos, porque tem situações muito diferentes na pós-graduação. Se por um lado o sucesso do sistema de pós-graduação é essa flexibilidade, por outro você tem uma variabilidade e muitas vezes essa variabilidade, no momento de fazer a avaliação, não é tão percebida, ou não é tão valorizada. Por quê? Um dos critérios para você montar as comissões de avaliação é: pessoas que tenham respeitabilidade no campo, que tenham vivência, que tenham experiência, ou seja: eles virão daqueles programas mais consolidados, de instituições mais fortes no campo da produção de pesquisa. Às vezes é difícil a pessoa se colocar no lugar de um curso pequeno, de uma instituição mais frágil, geograficamente mais distante dos centros de vida cultural e científica do país. Aí você começa a se dar conta dessas coisas, porque você começa a conhecer os programas por dentro, você visita, enfim...

Já como coordenadora eu comecei a perceber que esta avaliação estava muito parametrizada, muito estabelecida em cima de pontos de corte para todas as variáveis... Não é só para a produção, porque a gente

quando fala de avaliação o que mais salta é a questão da produção, porque de fato ela tem um peso muito central na avaliação. Mas as outras coisas também: quantos alunos por orientador, qual tempo de titulação, ou seja: tudo isso estava muito transformado em indicador aritmético, esquecendo que o número representa situações completamente diferentes. O mesmo número.

Como sou epidemiologista, trabalho muito com taxas, com números, então a gente é obrigada a ter essa percepção. E eu comecei a ver isso com relação à avaliação. Você muitas vezes se deixa levar pelo número, friamente, como se isso garantisse uma posição neutra do avaliador, como se fosse possível ser neutro ao avaliar. Avaliar é atribuir valor, não tem como escapar disso. Então as pessoas se escondem um pouco atrás dessa coisa, para dizer: “Olha, eu sou neutro, os números é que vão dar a posição”...

“A avaliação da Capes é comparativa. Ela não é essencialista. Quer dizer, não é uma avaliação que parte de essências, de princípios que deveriam estar sendo cumpridos. Ela parte de comparar entre os cursos que posição cada um tem”

Revista Adusp. Um positivismo feroz.

RBB. Feroz, nesse sentido, porque veja: por um lado, te protege, no sentido de que eu não conheço todos, então posso ser injusto; mas por outro, tira aquilo que a avaliação tem de essencial, que é compreender as situações que você está avaliando. Não é? Então, não é simples a situação. Claro que eu não tinha essa clareza na primeira avaliação que eu fiz. Na hora em que você vai para a segunda, e que estando no CTC você vê o que acontece em todas as áreas, e depois como diretora de Avaliação — aí por obrigação eu tinha que entender de todas as áreas, conhecer a diversidade — você vai vendo que as coisas vão se multiplicando em termos de dificuldades neste sentido que eu estou levantando. De uma avaliação muito amarrada numa ficha, a ficha muitas vezes não dá margem para avaliar aquilo que você está percebendo. A ficha é muito amarrada, ela tem pesos estabelecidos, o próprio sistema calcula e indica a nota. E muitas vezes não cabe naquele sistema tão amarrado a diversidade que você tem. Isso vale para tudo, vale para todos os aspectos da avaliação, não só para a questão da produção.

Na questão da produção a gente tem um complicador a mais, que é o fato de existirem métricas de avaliação da produção científica universalmente usadas, embora elas tenham sido muito pouco estudadas. Não é?



Porque tem uma grande discussão sobre os indicadores e as métricas que é argumentativa, baseada em raciocínio lógico, mas nem sempre a evidência empírica confirma esse raciocínio. E tem pouquíssimo estudo feito em cima do empírico. Então precisamos entender melhor o que estas métricas significam e em que medida elas podem ser úteis.

Por outro lado, nós temos um sistema de pós-graduação crescendo muito. Entre uma avaliação e outra, temos um crescimento de quase 25% no sistema, em quatro anos. Ou seja: é quase impossível a gente fazer a avaliação completamente qualitativa, porque o número de programas é enorme. Agora, talvez a saída esteja em combinar as duas coisas. Aquelas áreas que têm um número menor de programas em certa medida são melhor sucedidas em fazer isso. Elas usam os indicadores quantitativos, mas têm condições de analisar mais qualitativamente o conjunto de material que é oferecido para a avaliação, porque têm tempo para fazer isso. Elas têm, vamos dizer, 30

programas. Agora, uma área que tem 250 programas, e que tem a mesma semana para avaliar os 250... Tem uma coisa que é difícil as pessoas entenderem: por exemplo, essa área que tem 250 [programas] chama 80 consultores. Mas como é que você coordena 80 consultores? Então não resolve. Enfim, tem uma série de desafios na avaliação.

Voltando a falar dos indicadores de produção: eles acabam sendo usados em parte porque o mundo inteiro usa, e aí “bem, isso está consagrado”. Agora, o fato de todo mundo usar tem muito a ver com que é isso que está disponível. E mesmo [entre] as pessoas que são mais críticas ao uso dos indicadores quantitativos, poucas se dispõem a formular sugestões de como a gente poderia fazer de maneira objetiva uma avaliação qualitativa. Porque o risco que você corre, quando não tem isso objetivado, é deixar que as impressões predominem. E como a maioria das pessoas não conhece a totalidade dos cursos, a chance de você ser levado pelo prestígio que determinados

grupos têm, pelo fato de ter naquele programa alguém que é uma liderança no campo... Então a gente tem que trabalhar no sentido de desenvolver esses indicadores qualitativos mais objetiváveis, que não impliquem métrica, medir coisas, mas entender e compreender a essência dos programas. Esse é o desafio hoje, de enfrentar o chamado produtivismo.

Como é que isso aconteceu? Por quê? Vamos tentar entender. Até 1998, quando a Capes pela primeira vez formulou uma ficha única para todas as áreas, e passou a pôr uma ênfase muito forte na produção científica, mas na produção científica divulgada, seja livro, seja artigo, havia no Brasil, na pós-graduação principalmente, neste período de 1970 a 1990, uma produção que ia para a gaveta. As pessoas faziam projeto, desenvolviam pesquisas, mas consideravam que ao entregar o relatório de pesquisa estava terminado o trabalho. E a divulgação muitas vezes não se dava. Eu mesma tenho trabalhos desse período, de quando era mestra ainda ou estava fazendo doutorado, que foram para a gaveta. São pesquisas que poderiam ter dado uma contribuição importante para o sistema de saúde. Porque a gente não sentia a pressão. Você fez a pesquisa, entregou o relatório para o financiador, prestou suas contas, terminou. E ficava aquela... “a gente precisa escrever” aquela pesquisa. E a gente não escrevia. Porque não havia a pressão para isso, por um lado. Claro que sempre houve os intelectuais que divulgaram a sua pesquisa adequadamente. Mas havia muito disso entre os docentes. Até porque a gente trabalha demais, tem muitas tare-

fas, nunca tem gente suficiente para fazer tudo, isso vai ficando. Então a Capes de alguma forma disse: “Olha, a pesquisa só termina quando ela for publicada”. E a pós-graduação começou na avaliação a forçar isso. Num primeiro momento, OK. Positivo fazer isso. Mas o que acontece? A avaliação da Capes é comparativa. Ela não é essencialista. Não é assim: “Um curso 7 tem que ter essas virtudes para ser 7”. Não. Ele primeiro precisa ficar bem posicionado no *ranking* de todos os cursos da área, para depois eu olhar essas qualidades e dizer “ele pode ser 6, ou 7”.

Revista Adusp. A lógica da competição.

RBB. A lógica de ranqueamento. É essa. Está escrito no regulamento da pós. Quer dizer, ela não é uma avaliação que parte de essências, de princípios que deveriam estar sendo cumpridos. Ela parte de comparar entre os cursos que posição cada um tem. É claro que algumas áreas levam isso ao extremo, no sentido de realmente fazer uma coisa matemática, passar a régua: “Olha, vou dividir isso aqui em três pedaços”, e é assim. Tem áreas que não, que fazem uma mescla entre o que o indicador quantitativo está sugerindo e qual é o contexto desse programa. Vou dar um exemplo da minha área: a área de Saúde Coletiva. Quando eu fiz a minha primeira avaliação, tínhamos um programa de doutorado novo, instituído na [Universidade] Federal do Maranhão, que tinha um desempenho que não o colocaria entre os programas nota 4, nota mínima que um doutorado precisa ter. Nós fizemos toda uma argumentação perante o CTC de que, apesar

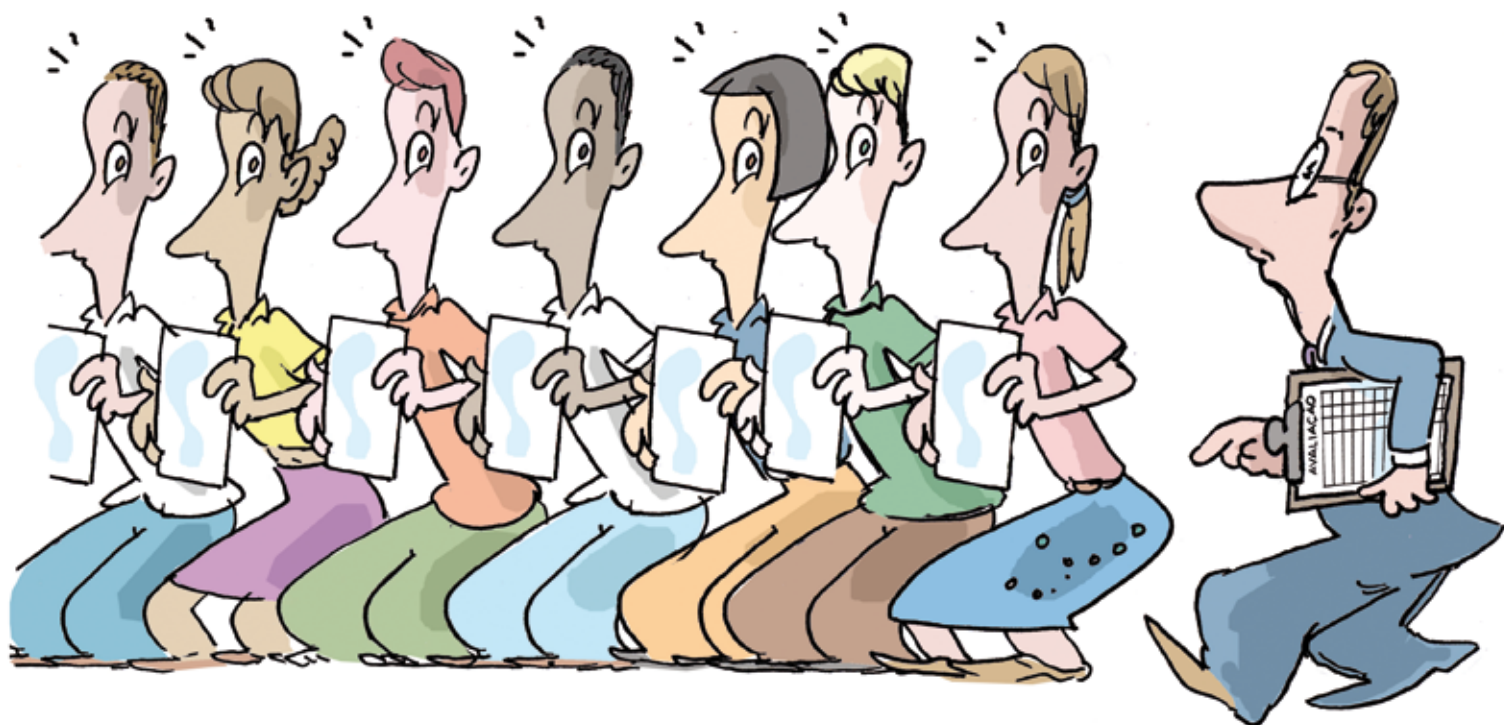
disso, nós estávamos considerando que era um curso novo, numa região onde não é fácil instalar um curso de Saúde Coletiva; que eles tinham liderança científica suficiente, pelo menos um dos docentes com liderança científica reconhecida no campo, para poder levar aquele curso para um outro lugar; e que a gente queria apostar naquele curso. Então é possível fazer isso, tanto é que muita gente diz: “É incoerência da análise”. Não, ao contrário: é onde a comissão realmente mostra os valores que ela está trazendo para a Pós, não é? Existe isso. Às vezes de fora as pessoas não conseguem enxergar que existe esse processo, mas muitas áreas trabalham assim. Usam o indicador quantitativo como uma referência inicial, mas durante a semana de avaliação a gente costuma ver o trabalho antes, e chegar lá já com as relatorias feitas, para poder ter uma discussão no conjunto. Aí você consegue olhar essas situações particulares e não cometer muitos erros. Outros grupos não fazem isso, ficam friamente naquilo que os indicadores quantitativos mostram. Eles defendem essa posição. A gente tem que respeitar, mas certamente você vai cometer mais enganos ao fazer isso, porque as situações são muito diversas, e a gente precisa levar isso em conta.

Revista Adusp. Eu vi a sua entrevista para o *Direto da Ciência*, a Sra. até menciona aquela imagem do “sarrafo”. A Sra. não acha que a própria lógica do sistema conduz a esse “sarrafo” sempre [mais alto]...?

RBB. Leva a isso. Sim! Porque veja, se é comparativo, e quando estou trabalhando o meu programa eu não

sei como é que os outros estão trabalhando, não adianta você dizer: “Olha, gente, a produção está OK, não precisa ir mais além”. O coordenador do programa pensa assim: “Bem, se eu não for e os outros forem, minha nota vai cair. Não vou ficar nem no lugar em que eu estou, quanto mais subir”. Então os coordenadores, ao selecionarem os orientadores, aplicam individualmente os valores que foram aplicados coletivamente na avaliação. Coletivamente, na área de Saúde Coletiva por exemplo, a média de produção dos programas está lá em 500 pontos, vamos dizer. É o conjunto. E o coordenador diz: “Ah, para ser orientador no meu programa você tem que fazer 500 pontos”. Porque é isso que precisa para tirar nota 5. Não é individual, o “500”. Mas é aplicado assim. Que é que isso vai produzir na próxima avaliação? Eu vou ter o que produziu 500, mas vou ter o que produziu 700. A média vai subindo. Então o próprio processo de avaliação, ao estar montado como está, e ao ser aplicado nos programas desse jeito, é uma máquina que não para de rodar. Uma máquina *maluca*, nesse sentido.

“A intenção era mudar, assim que terminou a Avaliação 2013. Que aconteceu? A Capes mudou de presidente três vezes, mudou o diretor de Avaliação três vezes. As coisas não têm continuidade. Não aconteceu”



Revista Adusp. Sim, mas é uma *máquina maluca* que se presta a um projeto de ciência, de universidade. A Sra. tem uma certa fé em algum modelo de avaliação. E esse modelo atual pretende introduzir uma lógica que é fabril, mercantil, contabilista. Queria ouvir sua opinião sobre isso.

RBB. Falando da experiência que eu tive com as direções da Capes, quando eu fui coordenadora, o professor Jorge Guimarães presidente e o professor Lívio [Amaral] diretor de Avaliação, e depois eu diretora de Avaliação, acho que já há na direção essa mesma percepção. Tanto é que se você pegar o relatório da Capes, da Diretoria de Avaliação, sobre a avaliação de 2013, já está escrito lá com todas as letras que esse modelo está exaurido. Que nós temos que substituir este modelo de avaliação por outro. Qual? Esta é a questão. Então o que nós discutimos agora,

2017? Você faz avaliação em 2013, referente ao período que termina em 2012. Você leva um ano fazendo a avaliação. Você vai poder rever o modelo no segundo ano do período que vai ser avaliado em seguida. Para esse período teria que ter feito isso no ano passado [2018], vai chegar 2020 e não mudou nada. Porque aí você quer mudar e as pessoas dizem: “Mas como vai mudar, agora que está no último ano da avaliação?”

Então ou você muda logo após a avaliação que terminou, ou você não consegue mudar. A intenção era mudar, assim que terminou a Avaliação 2013. Está lá escrito. O que aconteceu? A Capes mudou de presidente três vezes, mudou o diretor de Avaliação três vezes. Então as coisas não têm continuidade. Não aconteceu. Quando eu cheguei para fazer a Avaliação 2017 [o período avaliado passou de três para quatro anos], eu tive um ano

para preparar a avaliação, a gente já tentou fornecer para as comissões coisas um pouco diferentes. Mas não dava para mudar o modelo, porque todo mundo dizia: “Não, não podemos mudar agora que a avaliação é no ano que vem”, e num certo sentido é verdade, você não pode mudar completamente o jogo se os coordenadores e os programas se organizaram em função daquilo que estavam prevenido que seria a avaliação. Então nossa grande chance de mudar teria sido o ano passado. Só que não aconteceu. Portanto ela não vai acontecer para a próxima avaliação.

Revista Adusp. Mas nós entrevistamos um diretor de Avaliação anterior, Renato Janine, que mais tarde seria ministro, fizemos uma entrevista longa, ele já admitia algumas coisas, mas a Capes, no essencial, não muda. Do ponto de vista da Adusp, de outras entidades da área, são inú-

meras distorções que vêm sendo apontadas, por exemplo: disputa de poder entre áreas no CTC.

RBB. Agora não tem mais. Sabe por quê? Eu vivi as duas situações. O CTC pequeno e o CTC expandido. O professor Lívio, quando assumiu a diretoria da Capes, instituiu, embora o CTC continue formalmente sendo composto por representação, ele passou a fazer todas as reuniões de CTC com todos os coordenadores de área. Isso mudou completamente essa distribuição de poder entre as áreas. É nítido. Porque uma coisa é eu desclassificar uma área na ausência do representante daquela área. Outra coisa é eu fazer isso com ele sentado na minha frente. A partir do momento em que o CTC é ampliado, todas as áreas estão no CTC com direito a voz; voto quando se trata de deliberação não pode, porque não mudou no Estatuto. Eu batalhei bastante para que mudasse, mas não é do interesse do último presidente da Capes ampliar essa participação democrática. Eu custei muito para descobrir isso e foi um dos motivos de eu sair. Mas por mim já teria mudado no Estatuto, porque foi um avanço enorme o que o professor Lívio fez. É muito mais difícil você coordenar um colegiado com 50 membros — ou mais até, porque além dos 48 coordenadores de área tem representações da ANPG, dos pró-reitores de pesquisa e pós-graduação das universidades — do que um com 18 ou 20 pessoas. No entanto, o que você ganha em transparência, em acesso à informação, em menos manipulação... Veja, o que havia antes? Tudo era atribuído ao CTC. “O CTC não deixa”. Aí você vai ver, é o teu próprio representante de área

que dentro do conjunto de áreas quer manter uma posição de poder. Não é nem em relação às outras, é dentro da própria grande área, isso acontecia. Com o CTC ampliado, acabou. Hoje as pessoas são obrigadas a entender o que é a área do outro e a respeitar minimamente.

De alguma forma se recuou desta mudança, porque se voltou a ter o CTC só com os representantes, e para atenuar um pouco os problemas eles criaram uma instância que não é regimental, que é uma reunião prévia ao CTC entre os coordenadores por grande área. Por exemplo: a grande área da Saúde, que são nove coordenadores, se reúne e discute as coisas antes da reunião do CTC. Porque na reunião do CTC ela estará representada por três coordenadores, e não pelos nove. Então houve um recuo na última troca do CTC. Mas durante o período de 2013 até 2017 ele funcionou ampliado, com a participação de todos os coordenadores em todas as reuniões. Isto diminuiu muito o que de fato havia, e era um dos principais gargalos, que era esse poder, esse micropoder, vamos chamar assim, da representação em face às outras áreas que não estavam ali representadas. Então a circulação da informação era diferenciada, o poder que essas áreas tinham era maior em relação às outras, o desequilíbrio entre as áreas de Engenharias principalmente, Medicina também, essas áreas, vamos dizer assim, com um acúmulo de *capital político* no sistema de ciência e tecnologia brasileiro [e as demais], se reproduzia no CTC. Mas hoje isso está muito mais atenuado, e estava mais ainda quando o CTC era ampliado. Eu defendi

ferrenhamente essa manutenção do CTC ampliado porque acredito firmemente que nesse processo você melhorou todo mundo, porque as pessoas são obrigadas a se colocar no lugar do outro e a tentar entender o que é que o outro faz. E não achar que fazer ciência é só fazer o que eu faço na minha área.

“Nenhum país tem uma máquina como a Capes. Fui em várias reuniões. O México está montando sistema, Chile, Argentina, mas nenhum chega perto da Capes em quantidade de informação, processo de avaliação regular, periódico. Não tem”

Revista Adusp. Mas a senhora concorda que esse modelo se espalhou e se enraizou, teve filhotes? Por exemplo, na USP temos hoje um sistema de avaliação ultracentralizado, temos também a chamada Comissão Especial de Regimes de Trabalho (CERT), que é um sistema polialesco, que se baseia em coisas do tipo quantas citações você teve no Google Scholar.

RBB. Concordo. Na Europa você tem esse problema hoje com Bolonha e com a tentativa de uniformização. Nosso sistema se inspirou muito, em termos de “o que é que eu quero com a pós-graduação”, no modelo norte-americano. Só que os norte-americanos não fazem essas avaliações. A

avaliação lá é feita pelas avaliações das áreas, mais como referência de mercado. Quer dizer: esse curso é avaliado assim, você vai gastar o seu dinheiro, gaste sabendo o que você está comprando. Mas são as sociedades de especialistas, as sociedades científicas que fazem. O Estado não se mete, não tem nada a ver com nenhum dos processos de avaliação feitos nos Estados Unidos. Nenhum país tem uma máquina como a Capes. Fui em várias reuniões, discussões, nenhum lugar... Aqui na América Latina começa a haver. O México está montando um sistema, Chile, Argentina, mas nenhum chega perto do que a Capes faz em termos de quantidade de informação, processo de avaliação regular, periódico. Não tem. Na Europa você passa a ter isso por conta do financiamento para as instituições, mas também não é a avaliação de cursos, de programas. É a avaliação da unidade, da faculdade tal, do departamento tal, para repartir recursos. Então as situações são muito diferentes: o que está acontecendo lá do que acontece aqui. Mas tem consequências parecidas muitas vezes. Porque a mesma coisa está acontecendo na Europa agora: como é que a gente mede? Sempre foi um problema, como é que mede ciência? Não é de hoje que os pesquisadores e os próprios cientistas se perguntam: como é que mede, como é que eu sei que eu devo pôr mais dinheiro aqui, menos ali? Ou quanto de dinheiro eu preciso pôr aqui, quanto eu preciso pôr lá. É uma questão que está aí há muito tempo. Com os indicadores cientométricos, as pessoas dizem: “Por que a gente não usa esta coisa aqui, para de alguma maneira mensurar?” Acontece que a comparação

não é simples. Essa é a questão. Porque cada área tem uma tradição de produção de conhecimento diferente, então se você não levar essas coisas em consideração vai ter problemas inúmeros.

Neste sentido está tendo uma discussão na Capes, vai ter o novo Qualis no mês que vem [julho], a gente vai ver o que isso vai virar na prática, mas tem um grupo de trabalho discutindo como usar de maneira mais adequada esses indicadores, e não só pegando um valor absoluto do indicador, “a revista que atinge esse valor fica em tal lugar” e assim por diante. Muitas áreas já não fazem isso há muito tempo, mas o que passa para a comunidade são exatamente estas que batem o martelo assim. “Se o fator de impacto não for 2 vírgula ‘não sei quanto’, não vou classificar essa revista em tal lugar”. Mas isso é cada área que decide. Cada área tem critérios próprios para fazer isso. Não é o mesmo para todo mundo. A única coisa que é a mesma para todo mundo são aquelas regras de barreira que o CTC colocou para aumentar a capacidade de discriminação do Qualis, que estava completamente perdido. E foi o Renato [Janine] que fez essa análise, que levou à mudança.

Revista Adusp. Mas que acaba gerando um monte de distorções que se espalham. Um pesquisador do porte do Elisaldo Carlini queixou-se de que não conseguia publicar em português. “Eu preciso que as pessoas leiam isso, por que tenho de publicar em inglês?”

RBB. Tem uma série de questões. Agora, por outro lado você vai ter áreas que praticamente só publicam em português, e elas têm o Qualis do

mesmo jeito. Então as áreas de Humanidades e Ciências Sociais Aplicadas, tirando Economia, que tem esse mesmo padrão, as outras todas praticamente não têm publicação que não seja em português. Essa é uma outra discussão que se faz. O Qualis hoje não é comparável entre as áreas, de jeito nenhum. Claro que não é possível você fazer da forma burra, quer dizer: eu estabeleço os valores dos indicadores e todo mundo vai ter que aplicar isso. Isso não pode, porque é completamente equivocado. Agora, você precisaria descobrir uma forma de classificar que entendesse as diferenças de cada área. Porque são 800 mil artigos que foram produzidos pela pós-graduação brasileira, claro que com duplicatas, porque por exemplo se eu fizer um artigo com você, você é de um programa e eu sou de outro, ele vai entrar duas vezes. Ou eu faço com um colega. Isso a gente limpou na avaliação. Mas às vezes os coordenadores põem duas vezes, porque o dado sai direto do Lattes. Mas tem 800 mil artigos. Como é que vou fazer com 800 mil artigos?...

“O programa tem que ter diferentes perfis. Aquele cara que é um ótimo pesquisador, mas péssimo docente. E esse que gosta de ensinar vai produzir menos do que aquele que só pesquisa. Mas os dois são úteis para o programa”

Revista Adusp. E por que é que precisa fazer?

RBB. Não precisa com 800 mil artigos. Por isso a gente diz: tem que mudar a forma de avaliar. Então talvez não possa avaliar tudo. Talvez tenha que avaliar uma parte. Talvez tenha que levar em conta o que é mais relevante. Por isso a questão da autoavaliação do programa, que precisaria ter um peso maior. Que é como eles avaliam o que fazem de relevante. Não tem que avaliar cada pessoa, um erro enorme da pós-graduação é que não está avaliando programa na verdade: está avaliando soma de professores. E cada professor tem que ser igualzinho ao outro, no sistema atual, porque você tem as regrinhas: “Eu preciso ter três a cinco alunos”, “Eu preciso publicar xis”. Fica todo mundo igualzinho. Isso é péssimo para a produção de conhecimento. Para mim, hoje, o problema maior da avaliação está aí. Se tivesse que indicar qual é “o” problema, o problema é esse. Eu tenho que avaliar o programa, o programa tem que ter diversidade, eu preciso ter diferentes perfis. Eu tenho que ter aquele cara que é um ótimo pesquisador, mas é um péssimo docente. Que a gente sabe que existe. Então eu vou pôr os alunos para aprender com ele? E esse que gosta de ensinar vai produzir menos que aquele que faz só pesquisa. Mas os dois são úteis para o programa. Vamos pegar o exemplo do Carlini. Uma pessoa brilhante, uma pessoa que foi fundamental no país, formou gerações. Hoje ele está com mais idade. Eu não vou pedir para ele publicar o que publi-

ca um garoto de 25 ou de 30 anos. Agora, qual o programa que não se beneficiaria de o Carlini ser docente desse programa?

Hoje o sistema não permite isso. Porque o coordenador pega o resultado da última avaliação e gere o programa dele com aquilo. Ele não diz qual o melhor jeito de combinar os recursos que eu tenho aqui para fazer o melhor programa que eu consigo. Que era uma coisa que o Jorge [Guimarães] sempre dizia: “A regra da avaliação tem que ser essa. Eu tenho que fazer do melhor jeito possível, e a avaliação tem que ir lá e ver se eu estou fazendo”. Hoje inverteu completamente. Você tem os indicadores, os coordenadores tentam se basear naquilo para gerir os seus programas, e isso está resultando numa superpadronização dos perfis. Todo mundo tem que fazer do mesmo jeitinho. Uma pressão enorme por resultados de curto prazo, que é mortal para a ciência. Ninguém vai investir em nada novo, porque vai levar oito, dez anos para dar resultado. E os programas não estão dando a chance de você fazer isso.

Gerou uma série de disputas internas aos grupos, que não existiam antes de a avaliação ser desse jeito. Eu reconheço todos esses problemas. E eu reconheço por dentro, tendo vivenciado isso. Então estamos num momento muito desafiador de ter que fazer a mudança a qualquer custo. Mas eu não sinto que a gente tenha condições concretas, hoje, para que ela seja feita na intensidade que precisaria já para a próxima avaliação, para dar o recado. Para dizer: “Gente, vamos desmontar essa roda. Vamos fazer de outro jeito”.

“Hoje nós não conseguimos avaliar, a não ser pelo *paper* que o aluno eventualmente publica no trabalho de conclusão dele, se ele está de fato sendo bem formado. O *paper* é um *proxy* indireto dessa formação. Porque não sei quanto do que está ali foi ele que fez”

Eu participei como diretora de uma reunião que a área de Ciências da Computação faz anualmente, após a avaliação. Ela reúne todos os programas 6 e 7 num seminário, e convida cinco professores de Ciências da Computação de diferentes países, para assistirem a essas apresentações e fazerem sugestões. O que precisa mudar, o que precisa melhorar, quanto estamos distantes do que os melhores centros de pesquisa nessa área fazem, o que temos de diferente mas que é vantagem. Uma discussão bastante interessante. E me chamaram para o último dia, para assistir à apresentação dos professores. Impressionante: em três dias, a maioria deles acerta na mosca. Qual foi a grande mensagem? “OK, vocês encontraram uma maneira de avaliar a produção científica, mas e cadê a formação? Como é que eu sei que essas pessoas estão sendo bem formadas?”

Daniel Garcia



É isso. Hoje nós não conseguimos avaliar, a não ser pelo *paper* que o aluno eventualmente publica no trabalho de conclusão dele, se ele está de fato sendo bem formado. E o *paper* é um *proxy* indireto dessa formação. Porque eu não sei quanto do que está ali foi ele que fez, quanto é o trabalho como um todo do grupo no qual ele está inserido, qual é a autonomia que ele tem de produzir um outro trabalho daquele nível sozinho, em outra instituição, não é? Então nós estamos avaliando muito precariamente a formação. E este deveria ser o centro da mudança. Mas o que eu estou vendo das mudanças na ficha que está sendo proposta para a próxima avaliação é mais do mesmo. Sinto ter que dizer isto. Porque eu estava apostando numa mudança mais importante.

A gente queria não ter mais a ficha. Substituir a ficha por um roteiro de avaliação, mais livre, onde eu tivesse espaço e não ficasse amar-

rada nos ‘pesinhos’ e nas ponderações, para entender o que de fato um programa é. Ter uma clareza do que um programa jovem, numa instituição recente, é capaz de fazer. O que é que um programa com 40 anos de existência deveria estar fazendo. Quer dizer: minimamente, a avaliação precisaria dar conta dessas coisas. Eu tenho programas que estão há 20 anos fazendo exatamente o que faziam no ano da sua criação. Óbvio que isso não é bom. Agora, eu não posso querer que um curso que tem três, quatro anos de existência tenha o desempenho de um que tem 15 ou 20 [anos]. Porque é um processo, ninguém faz mágica. Ninguém começa do dia para a noite com um curso nota 7. Isso está mais do que demonstrado. As notas são correspondentes ao tempo de existência dos programas. Não é reta linear, mas tem muito a ver.

“Na verdade não existe um Qualis Livros. O que existe é um roteiro de análise. O número de livros de algumas áreas é impossível. O periódico é permanente. O livro é único. Ele não tem Qualis por isso: é um evento único”

Revista Adusp. Já sabemos de todos os problemas com o Qualis Periódicos. Qualis Livros não é uma aberração?

RBB. Na verdade não existe um Qualis Livros. O que existe é um roteiro de análise dos livros. Por quê? De novo, como para os artigos, o número de livros de algumas áreas é impossível. Impossível de fazer uma análise de conteúdo. Então se definiram algumas características e uma avaliação mais qualitativa para os estratos mais altos. Esse livro é inovador, traz realmente uma contribuição inovadora para o campo, ele inova na metodologia, ele inova no conteúdo, ele tem potencialidade de um impacto não só acadêmico mas social? Enfim, foram algumas coisas que a gente desenhou. De todo modo também está se tornando impossível fazer isso de maneira adequada. A Educação teve um número de livros na última avaliação que é impensável. No artigo, a gente avalia o periódico. O periódico é permanente, vamos dizer assim. O *New England* é o *New England*, há cinquenta anos. O livro é único. Ele não tem um Qualis por isso: é um evento único. Se fez no início a coisa de classificar editora, é um absurdo. Óbvio que não iria dar certo. Então se passou a olhar o livro, tanto é que as comissões recebem os livros fisicamente, ou pelo menos um PDF. Na última avaliação a Arquitetura fez tudo por meio eletrônico. Todo mundo tinha que escanear e mandar para o coordenador, para a comissão, eles abriram nos computadores, ninguém mandou o livro físico. Mas nas outras áreas se manda o livro físico, vai para uma biblioteca, essa biblioteca cataloga esses livros, fica com esses livros. Mas veja: não é um processo simples, de jeito nenhum, e tem que ter uma outra maneira.

O que a gente tinha começado a discutir? Quer usar a produção total, para dar uma olhada em como ela é? OK, mas a avaliação teria que ser feita com alguns exemplares. Amostra, não sorteada, o curso nos diz o que fez de melhor. “Quais as minhas melhores teses, nesse período que vocês vão avaliar?” Todo curso tem alunos medíocres, alunos médios e alunos excelentes. Está bom, eu vou te avaliar pelo que de melhor você fez. E é você que vai me dizer o que é o melhor, não sou eu. Está certo? O programa vai me dizer: “O que eu fiz de bacana nesse período que vocês vão avaliar”. Aí eu posso olhar os produtos. Posso montar uma comissão que vai ler, refletir, vai ler os livros, ler as teses. Claro que eu não vou avaliar de novo a tese, teve uma banca para fazer isso. Mas o que aquilo reflete do que o programa está fazendo.

Porque é isso que a gente deveria fazer na avaliação da pós. Em que medida essa produção me dá dicas de como esse programa está funcionando. É complicado? É complicado. Os coordenadores dos programas não vão gostar da atribuição, porque vai ter que haver uma discussão interna: o que eu seleciono como aquilo que vai ser o nosso cartão de visitas. Mas você avançaria, e desmontaria essa ideia de que todo mundo precisa fazer dez artigos, quinze artigos, vinte artigos. Não. Você precisa fazer um, muito bom. E para o aluno também. Uma coisa pequena, que me deixou satisfeita: não tem mais na avaliação o tempo médio de titulação. Custou muito para as pessoas entenderem que uma coisa é o tempo de bolsa, a instituição que dá a bolsa tem o direito de dizer: “Eu tenho dinheiro para

financiar 24 meses do mestrado e 48 do doutorado”. Agora, eu tenho aluno que em 48 meses não termina o doutorado dele. Então, que acaba acontecendo? Eu não deixo o aluno fazer uma investigação que eu sei que não será terminada em 48 meses. É isso que está acontecendo na prática. Mesmo para o mestrado, menos de 36 meses é uma loucura. Isso eu já fiquei contente, não está mais na ficha. Agora, precisa ver como cada instituição vai administrar isso. Porque uma coisa é bolsa. Outra coisa é o tempo que leva para um problema ser solucionado adequadamente. E no doutorado isso é fundamental. Você não pode banalizar, dar uma formação qualquer, porque esse cara não vai ter outra chance de se dedicar formalmente a uma formação, como ele tem quando está fazendo o doutorado.