

# **AVALIAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO: do ponto de vista de um nativo<sup>1</sup>**

**Claudia Fonseca, UFRGS**

**Resumo:** Neste artigo, a autora relata sua participação do novo sistema, instaurado há quatro anos pela CAPES, para a avaliação de programas de pós-graduação no Brasil. Depois de uma descrição das etapas da avaliação realizada em 1998, levanta algumas questões quanto à influência das novas tecnologias de avaliação sobre as atividades de pesquisa e ensino.

**Summary:** In this article, the author recounts her participation in the new audit system, implanted by CAPES four years ago, to evaluate post-graduate programs in Brazilian universities. After describing different phases of the 1998 evaluation process, she raises certain questions as to the long-term influence of the new audit technologies on teaching and research.

O sistema universitário brasileiro passa atualmente por diversas modificações profundas. Entre estas modificações, o novo sistema de avaliação de programas de pós-graduação introduzido em 1998 consta como uma de grande consequências. Considerando que, ainda depois de três anos e meio, o período de rodagem, pairam algumas dúvidas quanto ao processo, pensei contribuir para a reflexão com o seguinte documento sobre a avaliação de 1998. Trata-se, de certa forma, de um “relato nativo”, já que – enquanto representante de área -- participei, se não na concepção, pelo menos em todas as etapas da execução do processo. Ao mesmo tempo, tentei, na medida do possível, fazer uma descrição etnográfica do processo para transmitir algo do clima da avaliação. Com isso, minha intenção não é desvalidar o processo (a competência e seriedade dos colegas que se dedicam a essa tarefa ingrata são inquestionáveis), nem desmistificá-lo (quanto a isso, a perspicácia da comunidade científica antecede e vai além de qualquer comentário meu). Antes proponho pensar o sistema como o encontro particular entre pessoas e tecnologias, para relativizar noções que pudessem exagerar a influência de “viéses personalistas” daquelas, ou das maquinações deterministas destas.

---

<sup>1</sup> Apresentação discutida durante a XXI Reunião de ANPOCS, Caxambú 1998. Publicado em *Horizontes Antropológicos* 16: 261-275, 2002

## **Preâmbulo**

A avaliação dos programas de pós-graduação organizada pela CAPES em 1998 transcorreu em condições difíceis. O sistema de avaliação existia há mais de quinze anos. No entanto, naquele ano, foram introduzidas mudanças radicais, logo numa época de transtorno inaudito na política educacional. Tratava-se de uma conjuntura caracterizada por salários congelados, fundos limitados de pesquisa, corte de bolsa, sucateamento da infra-estrutura administrativa. Lembremos: o período de avaliação coincidiu com a greve de três meses nas universidades federais. Os acadêmicos chamados a integrar o Comitê Avaliador encontraram-se na situação esdrúxula de estar, por um lado, suspendendo todas as atividades docentes em protesto contra uma política educacional considerada pouco favorável à universidade pública e, por outro lado, aperfeiçoando um sistema de avaliação que servia para implementar essa mesma política - tudo em nome da qualidade acadêmica do sistema. Os detalhes desta situação merecem um estudo a parte. Aqui, levanto a questão para sublinhar a situação paradoxal do processo de avaliação a que me refiro. Coube a nós avaliadores a execução de uma tarefa supostamente rotineira apesar das bombas que caíam por todo lado<sup>2</sup>.

## **A “tarefa rotineira”**

Permitam-me falar dessa tarefa *como se fosse* rotineira. Antes de mais nada, quero destacar minha enorme admiração pelas pessoas que participaram do Comitê de Avaliação: Prof. Luiz Fernando Dias Duarte, Prof. Marcio Ferreira Silva, Prof. Rafael Menezes Bastos, Profa. Rita Laura Segato, e Profa. Maria Gabriela Martins d’Avila. A competência desta comissão é inquestionável - e o trabalho não foi pouco. Todos os membros

---

<sup>2</sup> Em 2001, três anos e uma nova avaliação (esta vez trienal) depois da experiência que descrevo nesse trabalho, encontramos-nos mais uma vez no meio de uma greve prolongada que reúne funcionários, professores e estudantes das principais universidades federais em busca de condições dignas de trabalho.

do Comitê tiveram que suspender importantes atividades, fazendo o possível e impossível para reorganizar agendas profissionais tremendamente carregadas, para se dedicar de corpo e alma a essa tarefa.

Creio que houve um espírito de seriedade que reflete o profissionalismo da área inteira - a Área de Antropologia. Fora os canais institucionalizados e transparentes, eu não recebi nenhum telefonema, nenhuma pressão para favorecer um programa ou outro. Em geral, a avaliação mostrou que os coordenadores de pós-graduação se dispuseram a entrar no jogo - gastando horas de seu tempo para preencher formulários do DATACAPES, e mostrando-se atentos a questões levantadas durante todas as discussões no Fórum de Coordenadores. Podemos dizer, portanto, que a avaliação foi fruto de um trabalho coletivo que reflete uma perspectiva honesta da nossa realidade acadêmica.

Cabe lembrar que a média dos conceitos da Área de Antropologia esteve acima da média geral - levando a CAPES, num primeiro momento, a questionar o rigor de nosso trabalho. Porém, depois da revisão de todas as áreas realizada por um comitê interdisciplinar, a CAPES reconheceu o alto nível da Área de Antropologia, endossando o trabalho do Comitê. Dentro dos limites do possível, o trabalho da avaliação foi bem feito. O processo, porém, não está livre de problemas. As bombas, afinal, continuam caindo e, neste artigo, quero levantar alguns elementos que podem nos ajudar a pensar como e onde nos situarmos na atual conjuntura.

### **Endireitando a pirâmide de conceitos**

Em 1998, a mudança no sistema foi radical. Não nos dávamos conta, antes de começar, quão radical. O sistema de avaliação corria num clima de relativa tranquilidade há mais de quinze anos. Fora algumas orientações gerais, cada comitê gozava de grande autonomia na formulação e aplicação de critérios de avaliação. Não eram critérios necessariamente quantificáveis, e, nesse sentido, não eram inteiramente “objetivos”. Mas não havia grande controvérsia, pois, por um processo de decantação, a grande maioria de programas subia progressivamente na escala até chegar ao conceito máximo. Em 1996, cinco dos oito programas da área de Antropologia estavam em nível A. O maior suspenso do processo era saber *se* e *quando* os outros programas iam subir de C para B, de B para A. O ranking confirmava “o

que já era sabido” - os programas mais antigos e consagrados ficavam em cima da hierarquia, e os mais novos esperavam pacientemente sua vez para entrar na mesma categoria. Do nosso ponto de vista, isto é, das áreas disciplinares, o sistema funcionava muito bem, obrigado.

Mas, do MEC e da CAPES, vinham sinais de descontentamento desde 1996. Diziam que este sistema, com tantos programas em conceito A, não avaliava mais nada. A distribuição ideal de conceitos, aquela pirâmide que coloca um pequeno número de programas em cima e a grande massa embaixo, tinha sido praticamente invertida. A mudança radical almejada pela CAPES tinha por objetivo endireitar a pirâmide, o que permitiria uma maior discriminação entre os programas. Essa política rompeu com quase vinte anos de tradição. O sistema anterior primava pelas relações cordiais; o princípio hierárquico regia um campo onde - mesmo se alguns (embaixo) resmungassem de vez em quando - havia uma solidariedade, uma identificação com o modelo. Era uma situação que, na opinião de algumas pessoas, levava à complacência e acomodação; na opinião de outras, estimulava a excelência da pós-graduação num clima de relativa cooperação e respeito mútuo.

Com as mudanças introduzidas em 1998, a CAPES visava uma ruptura e atingiu seu alvo. Quase a metade dos programas avaliados (considerando todas as áreas disciplinares) não acatou o conceito atribuído, entrando com recursos para receber uma reavaliação. Surgiu um clima tenso marcado pela desconfiança e a concorrência. Paradoxalmente, o novo sistema que propunha sanar dúvidas, criando clareza e transparência nos critérios de avaliação, engendrou boa dose de insegurança e confusão. Normal, disse a CAPES. Qualquer mudança importante produz esse tipo de efeito. Agora, cabe a cada área travar políticas sensatas para usar esta nova situação para o fortalecimento da área.

### **A participação da comunidade**

O princípio de avaliação não é nada novo no campo acadêmico. Usamos este princípio diariamente para julgar nossos alunos. E aprendemos dos historiadores que a noção de inspeção - por onde avaliadores externos julgam a qualidade do trabalho dos professores - é parte integrante do sistema moderno de educação. Aflora já no fim do século XVIII quando as universidades européias começam a trocar exames orais por exames escritos

como método mais confiável para medir resultados. Não havia portanto motivo para irmos contra a idéia de avaliação - desde que fosse um sistema pensado e operacionalizado por nós. Nesse sentido, a noção de “peer-group audit” - a avaliação por pares - torna-se uma noção chave do processo.

A comunidade acadêmica participou, efetivamente, no novo sistema da CAPES. Representantes eleitos por suas respectivas áreas compuseram comitês - com integrantes muitas vezes também indicados pela comunidade - que se familiarizaram desde o fim de 97 com as novas fichas. Essa “participação” não foi, porém, igual à de anos anteriores. Com as inovações, houve uma perda de autonomia do Comitê.

Em épocas anteriores, o Comitê reunia-se duas vezes, em dois períodos diferentes. Havia uma primeira reunião para os membros do Comitê discutirem e estabelecerem os critérios do “perfil A”. Já nessa reunião os membros do comitê recebiam todo o material que devia ser submetido à avaliação : pastas completas - todas as informações pertinentes reunidas num só documento sobre cada programa. Depois dessa primeira reunião, cada membro do Comitê levava para casa duas pastas (ou três, ou quatro), com o compromisso de estudar demoradamente as informações. Normalmente, tinha-se de seis semanas a dois meses para fazer circular as pastas por correio entre os membros do Comitê e para refletir sobre a adequação dos critérios abstratos de avaliação aos casos concretos. Chegava-se à segunda reunião do Comitê já com reflexões bem elaboradas sobre cada programa e seu lugar dentro da hierarquia de cursos. A reunião servia para comparar notas, chegar a um consenso quanto ao conceito final e, por último, traduzir esse consenso para os termos da ficha de avaliação.

Com as mudanças introduzidas nesse último biênio, o esquema de participação foi modificado. A “ficha de avaliação” que, em biênios anteriores, era usada apenas para traduzir nossa decisão em termos inteligíveis à CAPES, tornou-se, no novo processo, o beabá de nosso raciocínio. No final de 97, foi divulgada esta nova ficha - a mesma para todas as áreas acadêmicas - preparada por assessores designados pela CAPES. Os membros da comunidade acadêmica foram chamados a familiarizar-se com esse documento, aprendendo a pensar em termos de seis “quesitos”: proposta do programa, corpo docente, atividades de formação, linhas de pesquisa, defesas de tese e produção intelectual. Cada quesito tinha quatro itens que deviam ser avaliados numa escala de 1 a 4.

Em março, alguns membros do Comitê de Avaliação foram a Brasília receber instruções quanto à aplicação da ficha. Foi-nos atribuída uma série de tarefas: 1) determinar “pesos” para cada quesito e, (dentro de cada quesito) a cada item, permitindo, por exemplo, que o peso da “produção intelectual” fosse o dobro do de “atividades de formação”; 2) estabelecer os critérios numéricos para a atribuição de conceitos a cada item (por exemplo, o programa que, em média, produzia teses de mestrado em menos de 36 meses receberia “muito bom”, quem produzia em menos de 40 meses receberia “bom” e assim por diante). Juntas, essas duas operações deviam garantir uma fórmula matemática que produzisse o conceito final do programa, de 1 até 5. Cinco seria considerado equivalente ao antigo conceito A. Devíamos então 3) projetar critérios de excelência para os programas com nota 5 que, submetidos à nova etapa de avaliação, poderiam excepcionalmente ser classificados em 6 ou 7.

Tornou-se evidente já na primeira reunião que, não obstante as modificações que poderíamos introduzir, a ficha de avaliação refletia uma filosofia e um sistema de classificação que independiam de nossa vontade. Os membros do Comitê foram instruídos a fazer os necessários ajustes para adaptar o formulário à realidade da área, mas o parco número de dados disponíveis sobre essa realidade tornou essa tarefa quase impossível. Recebemos tabelas de dados agregados, expondo, para cada item, a média do conjunto de programas, mas ainda nessa fase, o comitê não dispunha de qualquer dado relativo aos programas individuais. Em outras palavras, não havia como visualizar curvas ou a distribuição de conceitos ENTRE os programas. Conscientes desses limites, procedemos, não sem reservas, à distribuição de tarefas. Em vez de levar para casa pastas com dados concretos sobre cada programa, levamos um exercício acadêmico - encontrar uma escala ideal para os seis quesitos da ficha .

Em maio, o Comitê se reuniu como um todo pela primeira e única vez com o intuito de aplicar a ficha e produzir um resultado final. A nossa curiosidade era grande, e em especial antecipávamos com prazer a idéia de finalmente ter em mãos dados qualitativos sobre os programas. Em vez disso, recebemos, no primeiro dia, um calhamaço com centenas de tabelas. A multiplicação de informações extravasava em muito os itens da ficha original. Devíamos, por exemplo, olhar o tempo de titulação desdobrado em média, moda e mediana; os titulados desdobrados em bolsistas e não

bolsistas; e os bolsistas desdobrados em “por até 12 meses”, 12-24 meses, e mais de 24 meses. Aprendemos a pensar em termos de NRD (sistema de classificação dos membros do corpo docente conforme tipo de vínculo e tempo de dedicação ao curso), e enfrentamos termos como “coeficiente curtose” (a não ser confundido com “coeficiente de variação”) - tudo, na crença de que iam nos dar a palavra final sobre a realidade de nossa área.

A partir do segundo dia, começamos a receber documentos mais completos voltados para um mesmo programa. Assim, chegou uma pilha de relatórios sobre as linhas de pesquisa - um para cada programa de pós-graduação; depois chegou uma pilha sobre teses e dissertações, e assim adiante. Em um desses documentos chamado “Programa” ainda recebemos um texto narrativo sobre os objetivos de cada programa, mas foi só no terceiro dia que finalmente chegou o “cadastro”, com os nomes dos docentes. Em outras palavras, a informação foi chegando em nossas mãos de maneira fragmentária. E fomos trabalhando conforme as circunstâncias - isto é, realizando uma comparação transversal entre os programas quanto a cada quesito.

Gastamos um tempo enorme tentando dominar aqueles números de forma a produzir um documento fiel à proposta da CAPES. Não tínhamos muita alternativa. Tentamos formular maneiras de levar em consideração a especificidade de cursos novos, de cursos que têm seleção de dois em dois anos, cursos que - por causa de exigências regimentais de sua reitoria - apresentam cálculos (por exemplo, de tempo médio de titulação) diferentes dos outros. Quando, finalmente, rodamos o programa, o resultado não deu muitas surpresas. A essas alturas, usando da margem de liberdade permitida pela CAPES, fizemos pequenos ajustes para que os programas correspondendo ao perfil de um curso “A” recebessem todos uma nota 5. Tínhamos então seis notas 5, três 4, e um 3.

Na última etapa da avaliação em que devíamos indicar se algum programa devia receber 6 ou 7, tivemos uma medida de autonomia equivalente à de biênios anteriores. Liberando-nos dos cálculos complicados, fixamo-nos em dois itens que seriam *o sine qua non* do aumento de conceito: a titulação de um número expressivo de doutores e uma boa produção intelectual pelo núcleo básico de docentes em veículos conceituados de circulação nacional e internacional. Dois dos programas foram eliminados por não terem produção suficiente de doutores, e dos quatro restantes apenas dois

registraram produção bibliográfica numa ampla gama de veículos nacionais e internacionais. A estes dois cursos foram atribuídos os conceitos “6” ou “7” conforme critérios precisos estabelecidos pelo Comitê.

A discussão quanto ao formulário a ser preenchido - critérios de avaliação e sua aplicação - é interminável. Foi uma discussão que nos ocupou durante longas horas. Mas, curiosamente, ninguém chegou a questionar seriamente os elementos básicos da nova avaliação que chegaram a nós já pre-fabricados: as metas, a linguagem, as modalidades. Na realidade, nas Ciências Humanas, apenas uma área resolveu manter a antiga filosofia de avaliação, atribuindo um conceito 6 a todos os que tinham recebido “A” na avaliação anterior (60% dos programas). Sua política não foi aceita pela CAPES e o Comitê foi obrigado a reunir-se de novo para adequar seu trabalho à política atual.

### **Sua própria racionalidade**

Marilyn Strathern, na sua análise da cultura da avaliação, levanta uma série de dúvidas quanto a esse procedimento<sup>3</sup>. Fazendo uma retrospectiva do sistema britânica, ela lembra que a primeiras avaliações ocorreram no início da década de 80, junto com os primeiros cortes orçamentários para as universidades, numa política thatcheriana de *mais alunos menos dinheiro*. “Era uma maneira de obrigar as universidades a dizerem para o governo o que cortar.”<sup>4</sup> Hoje, os administradores, aderindo a uma visão globalizada da realidade, emprestam modelos uns aos outros de forma a produzir - de Sidney a Londres, de Atenas a São Paulo - sistemas de avaliação surpreendentemente semelhantes.

O envolvimento da comunidade de pares se anuncia como elemento essencial do processo já que é uma marca do Estado liberal mostrar “que o poder governamental está disperso e o destino das pessoas reside em suas próprias mãos<sup>5</sup>”. Em princípio, haveria então uma possibilidade para cada

---

<sup>3</sup> Atendendo ao convite de Professora Mariza Corrêa, e financiada pela CAPES, passei dois dias em Campinas com Professora Marilyn Strathern que veio falar, entre outras coisas, da “cultura da avaliação”. As reflexões expressas neste artigo são fruto, em grande medida, deste encontro.

<sup>4</sup> As citações de M.Strathern, quando a fonte não é indicada de outra forma, são de sua palestra “A cultura da avaliação” proferida e gravada em vídeo no dia 17 de setembro, 1998 em Campinas.

<sup>5</sup> “It is a demonstration of liberal government to show that governance is disperse and people’s destinies are in their own hands.”



categoria profissional estabelecer seus próprios padrões de excelência, mas, curiosamente, o novo sistema parece carregar “sua própria trajetória, sua própria racionalidade, sua própria estética”.

Neste sistema, conforme Strathern, a balança pesa para o lado não necessariamente de quem mais trabalha, mas sim de quem produz provas de seu trabalho. Estamos na era de *information technology*. Na mentalidade produtivista, não há mais lugar, nas atividades de pesquisa, para “períodos vazios”, só de reflexão; os resultados devem ser imediatos e regulares. A *performance* que conta é aquela que melhor corresponde aos parâmetros da avaliação. Strathern descreve essa situação no quadro inglês: “As capacidades (sendo avaliadas) devem corresponder a certas expectativas, expectativas essas que espelham as capacidades necessárias para uma boa apresentação de dados a serem avaliados: clareza (antes de lógica), itemização (antes de conexões), marcadores [*bullet points*] (antes de parágrafos), e organização simplificada (antes da involução ou evolução do estilo argumentativo). Antes de tudo, abole-se ambigüidade, contradição ou hesitação” (1997: 315). O problema é que esse material produz uma “descrição etnográfica pobre”, revelando muito pouco sobre o verdadeiro funcionamento da organização.

Strathern refere-se a outra dimensão desse processo quando fala de uma premissa básica da cultura da avaliação: a descontextualização [*disembodiment*] de técnicas de ensino e procedimentos de pesquisa - em suma de todas as qualidades consideradas positivas para o sistema. Justamente por serem positivas, essas qualidades devem ser “exportáveis”, facilmente empacotadas para ser duplicadas em novos contextos. O êxito de um programa deve ser traduzido em termos que contribuam para a definição do padrão de excelência. A própria avaliação deve ser “objetiva” - isto é, passível de reprodução, com leves adaptações, em qualquer das áreas científicas.

Para comparar coisas complexas - programas de pós-graduação, por exemplo - é necessário reduzir o material a uma fórmula simplificada envolvendo poucas variáveis. Nesse processo, o contexto original tende a sumir. O “modelo ideal” (pois é com essa linguagem que o sistema de avaliação nos leva a raciocinar) é uno. Quem chega mais perto, recebe nota mais alta. Quanto mais singular o programa, quanto maior sua especificidade, maiores as dificuldades para avaliá-lo.

É irônico que, no quadro brasileiro, boa parte dos acadêmicos, especialmente nas ciências sociais, não tem grande afinidade com este tipo de análise. Na Antropologia, repudiamos em geral os “tropos” evolucionistas que distribuem as sociedades humanas numa mesma escala de mais e menos avançadas conforme critérios estanques; aprendemos a relativizar cada caso e analisá-lo em termos de contexto, da singularidade do evento. Este tipo de olhar é difícil de conciliar com as exigências da “cultura da avaliação”.

Outro paradoxo: Durante anos considerou-se que a melhor coisa para a Antropologia brasileira seria uma certa especialização dos programas. Para não haver simplesmente a multiplicação de um mesmo modelo, programas novos eram encorajados a desenvolver uma especialidade que faltava no cenário nacional. Hoje, essas especificidades contam pouco no processo de avaliação. Servem como brinde, depois de submetido cada programa à regra padronizada para todos. O tipo atual de cobrança impõe um padrão idêntico não somente para os programas mas também para os acadêmicos. Todos devem estar dando aula e fazendo pesquisa, mesmo quando participam ativamente da administração universitária. A divisão de tarefas que, conforme aptidões pessoais, permitiria uma concentração numa destas atividades à exclusão de outras, parece ser uma noção estranha a este modelo. Devemos lembrar que o DATACAPES é apenas um dos inúmeros formulários que os coordenadores de programa e professores preenchem durante o ano. Com o GED, acabamos de acrescentar mais um formulário à lista. A cada vez, aumenta a sensação de que os louros vão para quem sabe pôr o “x” no quadrado certo.

### **O boom de revistas**

Na avaliação CAPES, a maioria de comitês - incluindo o de Antropologia - deu destaque à categoria “produção intelectual”. Na lógica dos membros do Comitê um bom pesquisador deve mostrar resultados através de publicações em veículos de grande circulação - de preferência órgãos científicos de reconhecido mérito e ampla divulgação. Dessa forma o docente integra-se a debates acadêmicos de abrangência internacional. A questão que se coloca é como promover essa “circulação de saberes” sem engendrar a síndrome *publish or perish* - clima que privilegia a quantidade da produção em detrimento da qualidade.

Até 1995, a Antropologia brasileira contava apenas duas revistas nacionais: *A Revista da Antropologia* da USP e o *Anuário Antropológico* com sede na UNB. Em 95, foram lançadas *Mana* (Museu Nacional-UFRJ) e *Horizontes Antropológicos* (UFRGS). A Comissão de Avaliação reunida em 1998 já era confrontada a mais duas revistas - *Antropológicas* (UFF) e *Revista de Antropologia* (UFPe) e mais outra no horizonte, *Ilha* (UFSC).

Quando consideramos que existe atualmente um “boom” de revistas em outros departamentos das ciências sociais assim como nos institutos, o crescimento de periódicos científicos torna-se ainda mais impressionante. Parece que surge uma nova revista quase mensalmente. Não é difícil adivinhar o que provocou esta triplicação de revistas brasileiras nos últimos quatro anos. A maioria dos comitês de área estipulou que um curso “A” deve editar uma revista própria. Não esperamos que os critérios de avaliação sirvam para orientar políticas de melhoria nos programas de pós-graduação? A pergunta é : será que a produção intelectual aumentou em igual proporção?

A proliferação de revistas e publicações cria riscos que vão além da simples frustração de aspirantes a nota “A”. Como? Pensemos no modelo ideal de revista científica. Em razão de uma grande oferta de artigos, a revista consegue selecionar contribuições de qualidade, representando uma variedade de regiões e temas. Cria um banco de artigos publicáveis, acionando pareceristas para selecionar os melhores. A demora freqüentemente ligada ao trabalho dos pareceristas não atrasa a publicação graças à reserva de material.

A multiplicação exagerada de periódicos arrisca inverter esse processo. Cada programa andaria à cata de material para garantir a edição periódica de sua revista. Não haveria tempo nem para acumular um banco de reserva, nem para acionar pareceristas. Os contribuintes seriam na maioria da própria região assim como os leitores. Muitos destes autores, por fidelidade à revista do programa, deixariam de submeter artigos às revistas nacionais (e internacionais) de grande circulação e os debates intelectuais seriam assim regionalizados.

Não devemos ignorar a atual produção de coletâneas que, neste biênio, ultrapassou de longe a edição de livros por um só autor. Relacionando esta

onipresença de revistas e coletâneas à queixa dos editores de várias revistas tradicionais - que estão recebendo *poucos* manuscritos - temos causa de preocupação. Estaria a atual política de publicações propiciando a fecundação de idéias - o caráter aberto da troca intelectual - que queremos? É mais provável que ela tenha provocado uma dispersão de energias, uma paroquialização dos saberes e uma enorme perplexidade entre leitores potenciais que, diante do oceano de publicações, não têm idéia de onde iniciar seu programa de estudos.

Strathern descreve uma situação análoga na Inglaterra onde a explosão de publicações “teria corrompido o próprio sistema que a engendrou”. A proliferação frenética de revistas acadêmicas e livros teria ofuscado a distinção entre boa e má pesquisa. Desta forma, os avaliadores ingleses resolveram eliminar a *quantidade* de seus critérios para, a partir de 1996, apoiarem-se apenas na *qualidade* das publicações produzidas por cada docente<sup>6</sup>. Além de atentarem para o tipo de publicação (revistas científicas de maior ou menor prestígio), solicitam agora o envio de uma amostra da produção científica de cada docente (na média, um artigo por ano) e dedicam tempo (“o verão inteiro”) à leitura deste material.

## **O balanço - pros e contras**

Indagada sobre a reação de seus colegas britânicos ao sistema de avaliação, M. Strathern descreveu um quadro sem dúvida não muito diverso do que encontramos no meio acadêmico brasileiro. Evidentemente, não existe consenso. “Os colegas, na realidade, estão divididos. Há os complacentes que esperam que um dia a coisa toda vá simplesmente desaparecer. Há os novos administradores que a consideram salutar - pois ajuda a ‘enxugar’ o sistema. Há chefes de departamento que aproveitam o sistema para introduzir pequenas melhorias [*por exemplo, ao pressionar seus colegas a atualizarem seus programas e completar suas bibliografias de aula*]. E existe de fato uma série de maneiras pontuais de introduzir inovações

---

<sup>6</sup> Conforme M. Strathern, a avaliação de 1985 levou em consideração apenas a *quantidade* de publicações. As avaliações de 1989 e 1992 levaram em consideração tanto quantidade como qualidade de publicações.

genuínas. Não acredito, no entanto, que no balanço final as vantagens pesem mais do que os problemas.”

Estes problemas, que ameaçam a reprodução do sistema de ensino e pesquisa tal como o conhecemos hoje, tocam em particular nos seguintes itens:

- tempo: as idéias que produzimos até hoje foram fruto de um processo de reflexão e amadurecimento que exige tempo. O sistema de avaliação parece atropelar esse processo arriscando estancar a criatividade que é parte integrante da reflexão acadêmica.
- confiança: o sistema de avaliação não pode funcionar sem que haja confiança na competência e neutralidade dos colegas que avaliam. Paradoxalmente, a cultura da avaliação parece enfraquecer os laços de comunidade, minando a confiança mútua entre colegas.
- conjugação de pesquisa e ensino: embora existam hoje diversos lugares onde é possível realizar um trabalho intelectual interessante (Internet, companhias multinacionais, etc.), a universidade se destaca por ser o único lugar onde pesquisa anda junto com ensino, onde os esforços intelectuais podem desencadear uma reflexão que dura gerações. É uma combinação muito especial. O sistema de avaliação que coloca uma ênfase demasiada na quantidade (de alunos titulados ou de pesquisas publicadas) põe em perigo este arranjo particular.

## **Entre a tecnocracia e a reprodução do status quo**

Quando fui apresentada ao novo processo de avaliação, possuía um certo otimismo. Uma esperança que não foi inteiramente frustrada diz respeito ao lado “democrático” da avaliação : permitiria romper com o *status quo*, destacando pontos fortes e desnudando fraquezas em lugares inesperados. A própria Strathern cita como uma das vantagens da cultura de avaliação a possibilidade de revelar qualidades de colegas *junior* : “Eu acho que, geralmente, as regras ajudam as pessoas sem muito poder formal, ao passo que as detentoras de poder consideram estas mesmas regras como um incômodo”.

Eu tinha, no entanto, a expectativa de que o sistema de avaliação agisse em favor da comunidade acadêmica *como um todo* - e, a esse respeito, saí da experiência com mais dúvidas do que quando entrei. A idéia de que “parâmetros de excelência” fossem, de alguma forma contribuir para a alta qualidade da área já foi amplamente comentada nos parágrafos anteriores. Consideremos agora outra esperança – que os esforços de cada comitê permitissem à comunidade acadêmica exercer um mínimo de influência na política governamental de fomento. Os comitês nunca tiveram uma participação direta nesta política; pelo contrário, havia um esforço explícito de evitar a “contaminação” dos avaliadores por “este tipo de preocupação”. Os comitês eram chamados a produzir conceitos sem poder prever suas conseqüências. Mas esperavam exercer uma influência *indireta* ao fornecer uma avaliação criteriosa da área. Recentemente, o Prof. Maurício Barreto da UFBA sugeriu que esta esperança também seria em vão: apesar de todos os esforços da comunidade acadêmica para realizar uma avaliação coerente com a política da CAPES, esta instituição ignorou os conceitos na hora de distribuir bolsas. À base de cuidadosos cálculos, Professor Barreto conclui que “as concessões de bolsas para os cursos de mestrado foram definidas quase exclusivamente pelo tempo médio de titulação.”<sup>7</sup>

Em todo caso, apesar dos problemas, é provável que teremos que conviver durante muito tempo com a cultura de avaliação. Frente à expansão, nas últimas duas décadas, do sistema de ensino e do número de pesquisadores, seria praticamente impossível manter intocado o “antigo” sistema de pós-graduação e, por enquanto, ninguém imaginou uma maneira de coordenar pesquisa e ensino no atual contexto - além da cultura de avaliação.

Esta constatação não visa justificar uma atitude fatalista. Muito pelo contrário. Visa combater uma certa letargia que existe na comunidade. Devemos lembrar que - não obstante as análises foucaultianas - as instituições com as quais lidamos não são monolíticas. Existem entre os administradores correntes políticas diferentes, filosofias conflitantes. A CAPES - sensível a muitas das críticas levantadas aqui - está fazendo esforços para corrigir distorções e tornar o processo de avaliação mais ágil. A proposta de fazer “mini-avaliações” anualmente - um tipo de assessoria contínua que não desemboca em conceitos mas apenas em comentários

---

<sup>7</sup> Carta de 25 de março de 1999, de Prof. Maurício Barreto, dirigida ao Prof. Abílio Baeta Neves e distribuída pela Internet.

narrativos - é sem dúvida uma maneira de evitar o atropelo que ocorreu no último biênio. Se for levada a cabo, esta inovação deve permitir que os avaliadores acompanhem o processo de cada curso ao longo de três anos. Haverá tempo para ler exemplos da produção recente dos docentes e, em geral, desenvolver sobre cada curso uma reflexão amadurecida e compartilhada entre os membros do Comitê. Deveríamos procurar maneiras de introduzir outras modificações que minimizassem os estragos e ajustassem o sistema às necessidades da área.

Até recentemente, o processo de avaliação era visto com leve irritação - era uma inconveniência a ser esquecida o mais rápido possível. (Já não temos suficiente trabalho com nossas responsabilidades acadêmicas?) Quando muito, os programas reclamavam do conceito recebido mas sem iniciar nenhum movimento coletivo para articular uma crítica ou proferir sugestões substantivas. No vazio criado por esta vaga indiferença, as pessoas submetem-se ao sistema em vez de interagir com ele. Para evitar que as profecias pessimistas se realizem, será necessário reverter esta tendência canalizando uma porção maior de nossas energias para problemas que, no passado, nos pareciam “meramente burocráticos”.

Strathern, Marilyn. 1997. “‘Improving ratings’: audit in the British University system”. *European review* 5(3): 305-321.

.