

Os problemas do Enade



Por EWOUT TER HAAR & OTAVIANO HELENE*

Um exame padronizado simples como o Enade não é adequado para avaliar um sistema de ensino superior inovador e diverso, razão pela qual nenhum país no mundo usa metodologia similar

Está em pauta novamente, como aconteceu algumas vezes nos últimos vinte anos, a participação da USP no Exame Nacional de Desempenho dos Estudantes (Enade), a começar com os cursos da licenciatura. O acompanhamento sistemático e a avaliação dos cursos da Universidade são ações bem-vindas e necessárias. Certamente a USP deve se responsabilizar e criar mecanismos para esta tarefa. Porém, a eventual adesão ao Enade pela USP deve ser tratada com muita cautela, por três motivos principais.

Inicialmente, vamos argumentar que um exame padronizado simples como o Enade não é adequado para avaliar um sistema de ensino superior inovador e diverso, razão pela qual nenhum país no mundo usa metodologia similar. Um segundo motivo é que, ao participar do Enade, a USP submeteria seus programas de graduação desnecessariamente à regulação homogeneizante, desenhado para outros tipos de instituição, com o risco de estreitamento dos seus currículos. Finalmente, para os cursos da USP, o Enade iria gerar números altamente instáveis, inapropriados para fundamentar boas políticas educacionais.

É necessário saber quais são os objetivos de uma avaliação

É amplamente reconhecido, na comunidade de especialistas em avaliação educacional, que [a validade de um exame requer evidências a respeito do uso e interpretação dos seus resultados](#). O valor e validade do Enade devem ser julgados sempre em relação a como a USP vai usar e interpretar seus resultados. Na verdade, qualquer processo de avaliação deve ser precedido de um trabalho que defina a finalidade daquilo que se avalia.

Uma universidade, por exemplo, tem múltiplas funções, tais como formar quadros profissionais, fomentar o desenvolvimento social e cultural, contribuir para o avanço dos processos produtivos, produzir e se apropriar do conhecimento científico, entre outros. Além dessa múltipla finalidade, é necessário saber qual o objetivo de cada uma delas e até onde se quer chegar. Uma resposta simplória seria dizer que uma universidade tem todos aqueles propósitos e o objetivo final é o céu. Entretanto, tal tipo de resposta é irresponsável, pois se mais esforço for dirigido a uma finalidade, menos esforço sobrará para as demais.

Portanto, a pergunta que deve antecipar qualquer discussão sobre avaliação é saber o objetivo daquilo que se está avaliando. Há vários exemplos de casos em que essa pergunta ocorre antes da avaliação. Talvez as discussões referentes às grandes universidades do estado da Califórnia, EUA, possam ajudar a entender o processo. Um exemplo bastante conhecido e marcante foi [o planejamento do sistema de ensino superior daquele estado dos EUA](#), na década de 1960.

Atualmente, esse sistema tem três componentes: um sistema de *colleges*, com perto de dois milhões de estudantes, fornecendo formação profissional e cultural e servindo como um degrau para estudantes que queiram continuar o ensino superior em uma universidade; a Universidade do Estado da Califórnia, com mais do que 400 mil estudantes, que tem como principal finalidade a formação de profissionais, sem excluir, claro, a produção científica e cultural; e a Universidade da Califórnia, com perto de 300 mil estudantes, cujos principais objetivos incluem a produção de conhecimento científico, sem excluir a formação de profissionais.

Evidentemente, no caso deste exemplo, os critérios de avaliação de cada uma das partes desse sistema de ensino superior não podem ser iguais. O Enade, por outro lado, mesmo na sua versão nova, é um único exame, igual para todos os cursos de uma área, que busca caracterizar com apenas cem questões de múltipla escolha os quatro anos de formação dos concluintes de cursos de ensino superior, em somente uma ou duas dimensões. Perguntemos: qual é o papel dos cursos da USP na sociedade? O Enade ajuda a USP a avaliar se seus cursos contribuirão para “[formar profissionais líderes e cidadãos conscientes de seu papel social](#)”?

Theodore Porter, em seu [clássico livro sobre a história da construção de objetividade em políticas públicas](#), aponta para as vantagens políticas da criação de métricas supostamente objetivas para gestores (“é o algoritmo e a métrica que decidem”). Em sociedades como a brasileira, onde há um sentimento generalizado de que juízo e *expertise* humana são particularmente suspeitos e sujeitos à corrupção, o uso de provas e testes objetivos em processos avaliativos como concursos ou licitações é visto como indispensável. Por um lado, na educação, a cultura de avaliação externa e comparabilidade por meio de testes padronizados resultou em importantes avanços, como o Censo Escolar, o Saeb e os indicadores de gestão deles derivados, como o Ideb, que acompanham o ensino básico no Brasil inteiro. Mas é importante reconhecer os limites de validade de exames padronizados como o Enade.

A inadequação de exames padronizados para avaliar o ensino superior

O Brasil é o único país no mundo que usa uma prova objetiva para a dupla finalidade de avaliação e regulação do seu sistema de ensino superior. Quase nenhum outro país busca acompanhar a qualidade do seu sistema de ensino terciário por meio de um exame padronizado aplicado aos concluintes, no final dos seus cursos. As razões disso são expostas de forma cristalina [num relatório da OCDE](#), que a pedido do governo brasileiro investigou o seu sistema de “garantia de qualidade” dos cursos. O relatório, apesar de ser feito por uma organização conhecida pelo seu comprometimento com gestão de sistemas educacionais usando evidências quantitativas, é muito crítico ao atual sistema e mostra que os objetivos do Enade, medir a aprendizagem de egressos de cursos no ensino superior, são completamente inatingíveis, por três motivos principais.

Primeiro, ao contrário de uma avaliação no final do ensino básico, não é claro o que medir. Não há, e não deve haver, como ocorre no caso do ensino básico, um conjunto de competências e habilidades comuns que todos devem adquirir. É uma característica do ensino superior a diversidade de currículos e objetivos de aprendizagem ser muito grande. Um Enade que avalia somente as competências mais genéricas desvaloriza aqueles cursos capazes de trabalhar justamente aquelas competências mais especializadas que se espera de um curso de ensino superior. Mas qualquer avaliação de competências mais específicas seguiria necessariamente uma singular visão sobre os objetivos de um curso, em detrimento da pluralidade de visões valorizadas em estágios mais avançados do sistema educacional.

Um segundo desafio é técnico: como criar um exame relativamente curto que seja capaz de avaliar de forma confiável os conteúdos e competências adquiridas durante quatro anos de formação? Para fins de regulação, o novo Enade (por enquanto, para cursos da licenciatura) propõe usar a Teoria de Resposta ao Item (TRI), criando um exame de 45 questões de múltipla escolha para a dimensão Formação Geral e 60 para uma dimensão específica para cada área. Mesmo admitindo que competência em docência é um construto que pode ser capturado numa escala numérica de uma ou duas dimensões – uma preposição altamente problemática pelas razões expostas acima –, é ainda mais duvidosa que cem questões de múltipla escolha são capazes de avaliar de forma justa quatro anos de formação.

Terceiro, ao estabelecer um exame padronizado, que por razões práticas necessariamente é curto e focado num conjunto muito limitado de competências, há o risco enorme que os cursos venham a treinar seus estudantes para somente essas competências. Este estreitamento do currículo prejudicaria a possibilidade de inovações e a capacidade dos cursos de se adaptarem a mudanças e circunstâncias locais.

Todos esses motivos mencionados pelos especialistas da OCDE limitam gravemente as inferências justificadas que podem ser feitas a partir de testes padronizados no ensino superior. É por estes motivos que não há outros países no mundo que avaliem seus cursos no ensino superior por meio deste tipo de processo. Um projeto piloto promovido pela OCDE em 2013, a iniciativa AHELO, foi abandonado por ser julgado inviável. Na Europa e nos EUA, ao invés de depender de uma única forma de avaliação, sistemas de ensino superior são avaliados por meio de agências de acreditação que empregam formas amplas e holísticas de avaliação.

O novo Enade vai aumentar o número de questões e usar a TRI para ter comparabilidade entre anos de aplicação e melhorar a precisão. Como qualquer exame padronizado, o novo Enade tem o potencial de trazer evidências valiosas que podem ajudar a melhorar o rumo de cursos ou políticas públicas. Mas especialistas em avaliação educacional [avisa que um teste não necessariamente mede o que diz seu título](#) e que um pouco de mágica de equalização não deixa dois testes necessariamente equivalentes. A literatura sobre avaliação avisa especialmente sobre [a quase impossibilidade de usar um único exame para fins diagnósticos e fins regulatórios](#).

O funcionamento, validade, confiabilidade e interpretação do novo Enade devem ser estudados, antes de usar os números gerados por ele para fins de regulação e outras finalidades de alto valor. É muito preocupante o fato que o Inep, o órgão responsável pelo Enade, entre 2014 e 2021 calculou erroneamente o indicador mais importante derivado do Enade, de modo que [atribuía números essencialmente aleatórios a cursos](#). O erro foi corrigido finalmente em 2024, mas o fato de que nenhum dos interessados no processo de avaliação percebeu sua existência em todos esses anos coloca em dúvida o real papel regulatório dos indicadores do Inep, bem como sua capacidade de controlar a qualidade dos seus próprios indicadores.

Motivos específicos para a USP não usar o Enade

Para além das dificuldades gerais apontadas pela OCDE e especialistas em avaliação, no caso da USP há outros motivos para ser cético em relação à participação no Enade.

Primeiro, um ponto técnico. No Enade atual, descontado o efeito do Enem dos ingressantes, a variância das notas dos cursos é pequena: somente algo em torno de 10% da variância das notas dos alunos. Ou seja, a variação na nota Enade entre alunos de um determinado curso é muito maior do que a variação entre as médias dos cursos, e isso continuará a ser verdade para o novo Enade. Isso significa que qualquer indicador do curso derivado da nota Enade dos alunos deve ser calculado com muitos alunos participantes para ter confiabilidade estatística.

Acontece que, segundo o Censo de Educação Superior, somente cinco dos 34 cursos de licenciatura da USP têm mais do que 50 concluintes (média dos últimos cinco anos). Para a maioria dos cursos de licenciatura da USP (e também para outros de seus cursos), qualquer média ou outro indicador derivado do Enade será bastante instável e seria um erro adotar políticas neles baseadas.

Segundo, em princípio o exame poderia trazer evidências interessantes sobre os cursos da USP, se os resultados fossem usados para fins diagnósticos. Mas um argumento que está sendo circulado é que a participação do Enade poderia isentar os cursos da necessidade de renovação do reconhecimento pelo Conselho Estadual de Educação, mostrando claramente o desejo de que seus resultados sejam usados para fins de regulação.

Neste caso, haverá pressão quase inevitável para os cursos se guiarem pela matriz estreita do Enade ao invés de se

guiarem pelas Diretrizes Nacionais de Curso, que, propositalmente, deixam liberdade para os cursos desenharem seus programas de acordo com realidades locais. Da mesma forma, a aderência ao Enade enfraquecerá esforços da própria USP de implementar um programa de avaliação dos seus cursos mais amplo e adequado ao seu próprio contexto.

Conclusão

É improvável que a USP venha a aproveitar da participação do Enade; ao contrário, a submissão pela USP ao sistema federal de regulação, assim como é implementado hoje, implica grandes riscos para a qualidade dos nossos cursos. Haverá o risco de estreitamento dos currículos por causa dos efeitos secundários (*teaching to the test*) de ser avaliada por um exame simples demais, inadequado para a diversidade dos cursos da USP.

Ademais, aderindo a uma forma de regulação rasa, por indicadores instáveis e com pouca validade, a USP perde a chance de desenhar avaliações próprias, ricas e multifacetadas dos seus cursos. E para a maioria dos cursos da USP, com menos do que 50 concluintes por ano, os números gerados pelo Enade levarão a indicadores quase apenas aleatórios, instáveis demais para fundamentar políticas educacionais.

A USP deve se esforçar para implantar um sistema de avaliação própria dos seus cursos e usar sua *expertise* para pressionar os órgãos de controle de qualidade do ensino superior a usarem métodos mais válidos para fins de regulação. Para ser útil, um exame do tipo Enade apenas funcionaria num papel diagnóstico, para unidades de análise maiores do que cursos e de forma amostral, para evitar os efeitos colaterais da avaliação que destacamos.

***Ewout ter Haar** é professor do Instituto de Física da USP.

***Otaviano Helene** é professor sênior do Instituto de Física da USP.

Publicado originalmente no [Jornal da USP](#).

A Terra é Redonda existe graças aos nossos leitores e apoiadores.

Ajude-nos a manter esta ideia.

CONTRIBUA