

# Avaliando os resultados de aprendizado no ensino superior brasileiro

Este artigo foi publicado originalmente no periódico *Higher Education Management and Policy*, Vol. 24/2, OECD (2013), sob o título "Assessing higher education learning outcomes in Brazil"

## Introdução

No amplo contexto da expansão do ensino superior (ES) em todo o mundo, a avaliação da qualidade – dos sistemas nacionais, das instituições isoladamente consideradas e dos programas de graduação e pós-graduação – tornou-se um dos grandes temas nos debates envolvendo políticas para o setor terciário de ensino. Como muitos sistemas nacionais de ES estão passando por um intenso ciclo de expansão (algo que costuma envolver provedores de ensino não tradicionais que não são controlados diretamente por agências públicas, como empresas privadas e organizações com fins lucrativos, muitas delas empregando de maneira intensiva as tecnologias de ensino a distância), a avaliação e o controle de qualidade tornaram-se cada vez mais importantes. As demandas por responsabilidade e transparência vão além dos tradicionais procedimentos de credenciamento, que foram desenvolvidos por agências independentes ou governamentais, tipicamente envolvendo um relatório de autoavaliação por parte das instituições ou programas, além de uma visita oficial às instalações por parte de uma co-

## Renato Hyuda de Luna Pedrosa

Professor associado do Departamento de Política Científica e Tecnológica (Instituto de Geociências / Unicamp), tem se dedicado ao estudo de políticas relacionadas à Educação Superior, incluindo temas sobre transição do ensino médio à ES e equidade no acesso, incluindo ações afirmativas, sistemas de avaliação (Enem/Sinaes/Enade), o impacto da educação sobre o desenvolvimento econômico, qualidade da formação para as áreas CTEM (ciências, tecnologias, engenharias e matemática), governança e diversidade institucional da ES. Engenheiro Eletrônico pelo Instituto Tecnológico de Aeronáutica (1978), mestre em Matemática pela Unicamp (1981) na área de Teoria de Conjuntos e PhD em Matemática pela Universidade da Califórnia - Berkeley (1988), na área de Geometria Diferencial. É membro do Conselho Editorial da revista *Ensino Superior*.

## Eliana Martorano Amaral

Concluiu o mestrado (1992), doutorado em Tocoginecologia (1996) e livre-docência (2006) na Unicamp, onde é professora titular de Obstetrícia. Fez pós-doutorado na London School of Hygiene and Tropical Medicine, em 2001-2, em Pesquisa Aplicada à Saúde Sexual e Reprodutiva, com o tema da morbidade materna grave. Fez programa de especialização para pesquisa em ensino médico na Foundation for Advancement of Medical Education and Research na Filadélfia, em 2003-4. Coordenou a Comissão de Avaliação Profissional de 2008 a 2012 e é membro da Comissão de Ensino e Avaliação da Federação Brasileira de Ginecologia e Obstetrícia. É membro do Comitê Assessor em DST do Ministério da Saúde e consultora do Programa da Mulher da Secretaria de Saúde do Estado de São Paulo.

## Marcelo Knobel

Bacharel em Física (1989) e doutor em ciências (1992) pela Unicamp, tendo realizado pós-doutorado no Istituto Elettrotecnico Nazionale Galileo Ferraris, Turim, Itália, e no Instituto de Magnetismo Aplicado, Madri, Espanha. Professor titular do Instituto de Física Gleb Wataghin desde 1995, do Departamento de Física da Matéria Condensada, atuando na investigação experimental de materiais magnéticos nanoestruturados. Foi pró-reitor de graduação da Unicamp de 2009 a 2013, sendo responsável, entre outras ações, pela implantação do Programa Interdisciplinar de Educação Superior (ProfIS), que alia inclusão social com formação geral. Membro (notório saber) da Comissão Nacional de Avaliação da Educação Superior (Conaes). Editor-chefe da Revista *Ciência & Cultura* da Sociedade Brasileira para o Progresso da Ciência (SBPC) e membro do Conselho Editorial da revista *Ensino Superior*. Em 2010, foi laureado Comendador da Ordem do Mérito Científico pela Presidência da República.

missão de especialistas. A nova tendência é utilizar os resultados de aprendizagem dos estudantes para avaliar os programas de ensino superior e as instituições que oferecem esse serviço.

Tal aspecto do controle de qualidade no ES atraiu atenção cada vez maior tanto por parte dos especialistas quanto do público, incluindo responsáveis pelas políticas públicas e outros envolvidos. Como exemplo, um estudo recente (Arum e Roksa, 2011) apresentando um panorama desanimador da proficiência dos estudantes dos Estados Unidos durante os anos passados na faculdade – em áreas como leitura, pensamento crítico, construção de argumentos e outras habilidades – foi amplamente debatido pela mídia, pelos responsáveis pelas políticas públicas e pelo público em geral. Sua metodologia incluiu o uso de informações a respeito do desempenho dos estudantes num teste padronizado conhecido como Collegiate Learning Assessment (CLA). A validade do CLA enquanto instrumento para avaliar os resultados de aprendizagem foi amplamente debatida (Klein et al., 2007; Banta, 2006 e 2007; Pike, 2006).

Nos anos mais recentes, apelos em favor da avaliação dos resultados de

aprendizado tornaram-se mais frequentes, tanto por parte de agentes públicos quanto de analistas na mídia (Brooks, 2012). Já em 2006, o relatório da Comissão Americana para o Futuro do Ensino Superior (também conhecida como Comissão Spellings) pediu que o CLA fosse usado para avaliar os resultados de aprendizado com base no “valor agregado” (Departamento de Educação dos EUA, 2006:24), e que os dados fossem usados para propósitos de credenciamento (Departamento de Educação dos EUA, 2006:25), além de revelados ao público.

Em 2009, a União Europeia (UE) pediu à divisão de ensino superior da OCDE (Gestão Institucional no Ensino Superior, IMHE) que desenvolvesse o projeto Avaliação dos Resultados de Aprendizado do Ensino Superior (AHELO) (AHELO, 2011, Coates e Richardson, 2011). O projeto concluiu a fase piloto no final de 2012 e os relatórios estão agora disponíveis na rede<sup>[1]</sup>. O veículo de avaliação dos resultados de aprendizado inclui componentes do ensino geral e do ensino por tema.

Mesmo antes dessas iniciativas, alguns países já estavam experimentando com sistemas para avaliar os resultados de aprendizado no ES (Nusche, 2008). O Brasil foi talvez o primeiro a desenvolver um abrangente sistema de avaliação do ES, incluindo um componente de avaliação dos resultados de aprendizado, tanto para as instituições públicas quanto para as particulares: o Sistema Nacional de Avaliação do Ensino Superior (Sinaes). Entre seus componentes, inclui um teste para avaliar o resultado de aprendizagem (obrigatório para os estudantes), o Exame Nacional de Desempenho de Estudantes (Enade), com um formato muito mais próximo do espírito proposto pelo projeto AHELO. O propósito deste artigo

**O Brasil foi talvez o primeiro a desenvolver um abrangente sistema de avaliação do ES, incluindo um componente de avaliação dos resultados de aprendizado, tanto para as instituições públicas quanto para as particulares: o Sistema Nacional de Avaliação do Ensino Superior (Sinaes).**

1 <http://www.oecd.org/edu/skills-beyond-school/ahelodocuments.htm>

é apresentar uma visão geral do sistema brasileiro, debater questões de validade para o uso proposto para os resultados e outros ligados à natureza gigantesca e diversificada do sistema que é objeto desta análise. Também é debatido o envolvimento dos estudantes, a percepção deles em relação ao sistema e sua atitude em relação à participação.

### COMO SINAES E ENADE FUNCIONAM PARA GERAR INFORMAÇÕES DE AVALIAÇÃO

A lei federal criou o Sinaes e seu componente de exame, Enade, em 2004, sob os auspícios do Ministério da Educação. A aprovação da lei se seguiu a um acalorado debate acadêmico e político envolvendo um sistema anterior que também contava com um elemento de exame (Verhine et al., 2006). A legislação também criou a Comissão Nacional de Avaliação do Ensino Superior (Conaes), composto por funcionários do ministério e representantes de várias organizações de ES, incluindo instituições locais, estaduais e federais, para supervisionar todas as atividades do Sinaes/Enade e todas as decisões ligadas a mudanças no sistema. O Instituto Nacional de Estudos e Pesquisas Educacionais (Inep) é o órgão ministerial responsável pelo desenvolvimento e aplicação do Enade.

Embora o foco deste artigo seja o componente de avaliação dos resultados de aprendizado (Enade) do Sinaes, primeiro vamos delinear brevemente o sistema geral de avaliação para inserir o exame no devido contexto. Uma descrição mais detalhada do sistema completo pode ser encontrada em Verhine et al. 2006 ou no relatório do Inep a respeito do sistema (Inep, 2009).

Desde a sua concepção, o Sinaes teve como base três componentes principais: a avaliação da proficiência estudantil em

Um problema do **Enade** é que todas **as notas** têm como base um modelo de “normas de referência”, o que significa que **são relativas** e dependem de dados de todas as instituições/cursos, e **não têm como base um nível esperado de desempenho**. Não fica claro, portanto, se uma nota **quatro ou cinco** para a turma de formandos no Enade significaria que eles **de fato atingiram um nível de proficiência esperada**.

graduação, a avaliação institucional e a avaliação dos programas de ensino. O Enade, aplicado aos estudantes de primeiro<sup>[2]</sup> e último ano, avalia a proficiência dos estudantes em temas determinados pelas Diretrizes Curriculares Nacionais para os programas de graduação. A avaliação de programas e instituições tem como base dados reunidos pelo Ministério da Educação e em autoavaliações realizadas pelas instituições. Os programas de pós-graduação são avaliados pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), sistema que tem um prolongado histórico de sucesso no ES brasileiro, desde meados dos anos 1970 (Balbachevsky, 2004; Balbachevsky e Schwartzman, 2010).

Os resultados do Enade são derivados a partir de numerosas variáveis de entrada, incluindo as notas das turmas de formandos, um componente de valor

<sup>2</sup> A partir da edição de 2011 do Enade apenas os estudantes de último ano participaram do exame.

agregado que usa os resultados dos estudantes de primeiro ano e as informações a respeito da infraestrutura (instalações, bibliotecas etc.), corpo docente e programação dos cursos. Estes foram obtidos a partir do banco de dados institucional, e então combinados numa nota final, o Conceito Preliminar de Curso (CPC) de cada programa de graduação. As variáveis de entrada são ponderadas de acordo com a seguinte estrutura de peso: 40% da nota do Enade para os formandos, 30% do índice de valor agregado e 30% dos dados das instituições e programas. O CPC é atribuído numa escala de um a cinco, sendo cinco a nota mais alta.

Combinação das notas do CPC (graduação) e da Capes (pós-graduação) para os programas de ensino de uma determinada instituição, o Índice Geral de Cursos (IGC) dessa universidade ou faculdade é computado também numa escala de um a cinco. Por fim, a instituição produz um relatório de autoavaliação que, somado ao IGC, resulta no Conceito Institucional, envolvendo vários aspectos da atividade em ES, que são definidos pelo Conaes e determinados pela legislação.

Mais alguns detalhes a respeito do Enade: ele usa uma estrutura de censo

obrigatório, o que significa que todos os estudantes do último ano de cada curso avaliado são obrigados a fazer o exame (até a edição de 2010, todos os estudantes de primeiro ano também eram obrigados a fazer o exame)<sup>3</sup>. Essas características do sistema têm muitas consequências, em termos da confiança geral nos resultados do sistema de avaliação, incluindo a questão do envolvimento dos estudantes.

Um segundo aspecto é o de que todas as notas têm como base um modelo de “normas de referência”, o que significa que as notas são relativas e dependem de dados de todas as instituições/cursos, e não têm como base um nível esperado de desempenho. Assim, não fica claro se uma nota quatro ou cinco para a turma de formandos no Enade significaria que eles de fato atingiram um nível de proficiência esperada. Da mesma maneira, conceitos preliminares de curso (CPCs) também emanam de uma análise relativa, sem representar o nível de excelência de um programa em si. Uma alternativa seria uma abordagem com “critérios de referência”, exigindo uma escala de proficiência e

3 Quando o sistema foi adotado pela primeira vez, uma amostragem dos estudantes de cada curso participou do Enade. A partir de 2010, foi adotada a estrutura de censo.

Foram **poucas as evidências documentadas** e apresentadas ao público geral do sistema (por exemplo, estudantes, universidades, políticos etc.) para as decisões tomadas. A **validade** de qualquer sistema de avaliação do ensino depende muito da **apresentação de argumentos convincentes** para as decisões tomadas, mostrando que os resultados foram **bem documentados e justificados**; algo que os especialistas pedem desde a concepção do Sinaes.

## O Ministério da Educação emitiu alguns relatórios técnicos explicando a metodologia empregada pelo Sinaes e o uso das notas do Enade, mas, até o momento, nenhuma análise abrangente de validade dos usos dos resultados do sistema foi desenvolvida.

interpretação a ser dada, e uma definição do que seria um desempenho adequado. Entretanto, é importante observar que não está claro se tal escala de proficiência e critérios adequados poderia ser estabelecida para os cursos do ensino superior (Verhine et al., 2006). A maioria dos sistemas existentes de avaliação do resultado do aprendizado no ES usa um sistema de normas de referência (Nusche, 2008).

Dado o tamanho das populações envolvidas e a grande variedade de cursos avaliados, o sistema todo opera num ciclo de rotação de três anos: todos os cursos são agrupados em três áreas amplas, e cada grupo de programas é avaliado a cada três anos. Um ciclo completo de avaliação é concluído a cada três anos, quando todos os programas de cada instituição tiverem passado por todo o sistema de avaliação. Isso traz algumas questões envolvendo a validade dos resultados para alguns programas, já que o número de estudantes numa turma de formandos pode ser relativamente pequeno, o que significa que pode demorar bastante, talvez dois ou três ciclos, até que algo realmente significativo possa ser dito a respeito de um curso desses em termos de resultados de aprendizado.

### ENADE EM DETALHES

A edição de 2010 do Enade é usada como principal referência para descrever o desenvolvimento do exame; e ela é suficiente, pois a estrutura geral do exame se

mantém consistente nas diferentes áreas e temas no decorrer do período transcorrido desde o início da aplicação do exame.

O Enade tem duas partes: componentes de ensino geral e de áreas específicas do conhecimento. A porção do exame ligada ao ensino geral consiste em 10 itens, 8 questões de múltipla escolha (ME) e dois itens de tema livre (ensaios breves). A porção relacionada às áreas específicas consiste em 27 MEs e 3 itens de tema livre. O peso na nota final é de 25% para o ensino geral e 75% para as áreas específicas. Os estudantes têm quatro horas para concluir o teste.

A porção ligada ao ensino geral é comum a todos os cursos que participam num determinado ano e seu conteúdo não está relacionado ao curso frequentado por um estudante. Em 2010, por exemplo, os cursos de ciências da saúde e ciências agropecuárias foram avaliados. Em 2011, todos os cursos de engenharia, ciência básica e treinamento de professores foram avaliados. Entretanto, diferentemente do CLA, criado para avaliar “capacidades básicas de escrita e pensamento crítico”, o componente de avaliação geral do Enade avalia o “conteúdo do ensino geral”; em outras palavras, o conteúdo do exame está ligado a aspectos culturais e sociais da sociedade contemporânea. Essa abordagem traz problemas para a interpretação dos resultados, pois supõe uma relação entre os aspectos avaliados e o ensino universitário no Brasil. Além disso, como

a lei brasileira não exige que os cursos de ES incluam conteúdo do “ensino geral”, a variação nos conhecimentos não necessariamente ligados ao ES entre os alunos distorce as notas. Parece que os criadores do exame imaginaram que uma “boa” faculdade ofereceria esse tipo de ensino aos alunos. Por outro lado, pode ser um indício da direção que o Ministério da Educação gostaria de dar para a evolução do ensino superior brasileiro.

Pode-se supor que um teste para conhecimentos tão gerais teria resultados bastante correlacionados ao ensino recebido pelos estudantes ainda no ensino médio<sup>4</sup>. Uma análise rápida das tabelas de resultados do Enade mostra que as notas dos estudantes de Medicina do primeiro ano (curso altamente seletivo) são mais altas do que as dos estudantes de primeiro ano de Enfermagem (curso menos seletivo). Portanto, é provável que uma análise mais cuidadosa revele que os resultados do teste de conhecimen-

tos gerais estejam fortemente ligados ao contexto socioeconômico ou à qualidade do ensino recebido pelos alunos antes de ingressarem no ES.

A porção do Enade que avalia conhecimentos específicos é mais parecida com a do AHELO. Inclui a avaliação de áreas básicas num curso de graduação, avançando gradualmente dos temas mais básicos até os mais especializados, de acordo com aqueles que são em geral considerados os temas mais relevantes no ensino de graduação para a área específica. Os 27 itens de ME no teste envolvem questões e problemas a respeito de diferentes aspectos do conteúdo do curso. O mesmo pode ser dito dos três itens na forma de ensaios escritos, que costumam exigir mais trabalho do estudante para se chegar a uma resposta adequada. Do ponto de vista qualitativo, essa parte do exame pode ser considerada adequada, embora não tenham sido realizados estudos de validação para testar essa suposição.

4 Souza (2006) apresenta evidências que apontam nessa direção, mesmo para a parte específica do teste, para os cursos de Contabilidade.

#### **USOS DO SINAES/ENADE: RESPONSABILIDADE, TRANSPARÊNCIA E APOIO POLÍTICO**

O Ministério da Educação usa as várias notas e relatórios produzidos durante os três anos de um ciclo do Sinaes para propósitos de credenciamento nos níveis institucional e dos cursos. A proposta original do Sinaes (Inep, 2009) enfatizava mais a avaliação de instituições e processos do que a nota final ou a nota do Enade. Entretanto, a divulgação das notas e dos resultados de credenciamento por parte da mídia alterou gradualmente o foco, voltando-o para tais aspectos, em especial depois que diferentes publicações impressas começaram a produzir classificações com base nos resultados.

Todos os anos há instituições que são colocadas em observação quando os cursos apresentam desempenho fraco nos

**Além de tratar da mesma maneira cursos muito diferentes, como treinamento de docentes e Engenharia, os mesmos princípios também são usados para avaliar todas as instituições, desconsiderando características acadêmicas específicas ligadas à sua missão, como a ênfase na pesquisa ou no ensino.**

dois ciclos mais recentes. Então, o Ministério impõe cortes no número de vagas para matrícula como consequência do período de observação. Entre 2007 e 2011, por exemplo, o Ministério cortou 1.114 vagas em cursos de Medicina, sendo 514 delas somente em 2011, após a rodada de 2010 do Sinaes. Nos cursos de Direito<sup>5</sup>, nada menos do que 34 mil vagas foram cortadas desde 2007. Muitos cursos foram descredenciados e houve um caso em que uma instituição inteira foi fechada, com seus estudantes transferidos para universidades vizinhas.

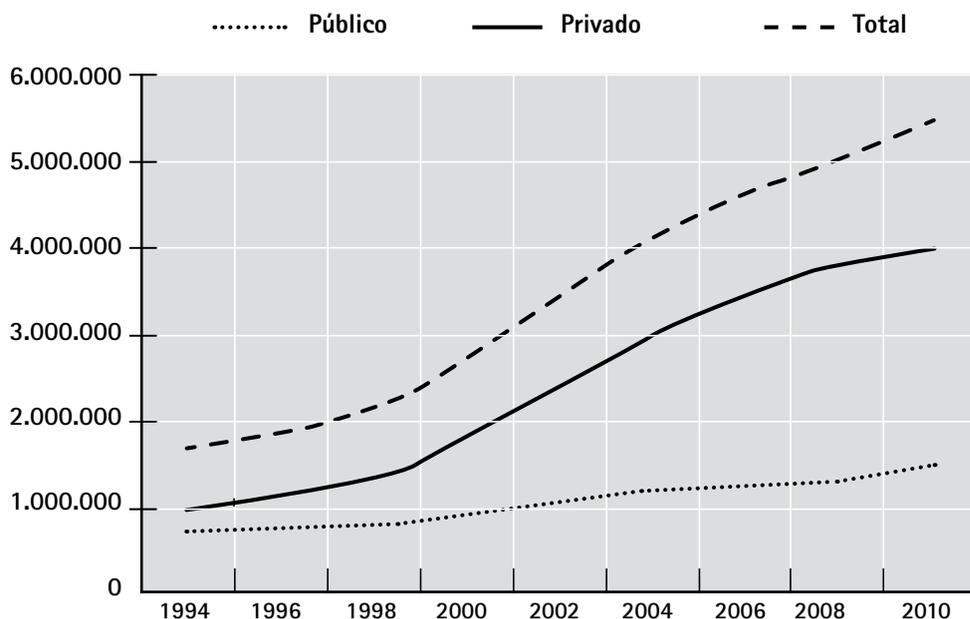
Essas medidas ajudaram o Enade, em particular, e o sistema como um todo a conquistarem apoio público e político.

5 O Direito é um dos cursos mais procurados do sistema brasileiro, chegando a quase 700 mil matriculados em 2010, o que corresponde a quase 13% do total de matrículas de graduação no Brasil naquele ano.

co. Tipicamente, as instituições que têm vagas cortadas ou cursos fechados são faculdades e universidades particulares com fins lucrativos. Parte do raciocínio apresentado por aqueles que apoiam o sistema está ligada à aparente necessidade de regulamentação do setor privado, em rápido crescimento. Esse setor quadruplicou de tamanho desde 1994, como podemos ver na **Figura 1**. O número de matrículas no setor privado apresentou aumento de 400%, passando de pouco menos de 1 milhão para quase 4 milhões de estudantes, correspondendo a 73% do total de 5,4 milhões de estudantes matriculados em 2010 (um aumento em relação a 1994, quando esta parcela era de 58%).

O perfil do setor privado também mudou drasticamente: em 2010, dois

Figura 1. Matrículas de graduação em cursos regulares presenciais  
BRASIL, 1994-2010



Fonte: Inep (1994-2010), "Sinopses Estatísticas da Educação Superior – Graduação", Ministério da Educação, disponível em <http://portal.inep.gov.br/superior-censosuperior-sinopse>.

Da mesma forma, **universidades tecnológicas** (que oferecem cursos vocacionais de 3 anos), subsistema em expansão, são tratadas **de maneira semelhante às universidades**, com o prevalectimento óbvio dos critérios de qualidade de tipo universitário, como **número de docentes com diploma de doutorado e indicadores de pesquisa**.

terços dos estudantes estavam matriculados em instituições de fins lucrativos, enquanto em 1994 mais de dois terços dos estudantes estavam matriculados em instituições de ES sem fins lucrativos<sup>[6]</sup>.

Todos os dados produzidos pelo Ministério usando o Enade são divulgados ao público. Embora condizente com as metas de transparência, isso também facilita o trabalho da mídia interessada em desenvolver “tabelas» e classificações de instituições/cursos, com todos os problemas envolvidos na manipulação ou ocultamento de informações para atender a propósitos específicos ou por falta de conhecimento. Apenas uma das notas do Enade é usada no desenvolvimento de classificações, por exemplo, dependendo das intenções do responsável pela clas-

6 Esses números referem-se a estudantes matriculados em cursos presenciais de graduação; os cursos a distância também passam por um período de crescimento acelerado, e já há quase 1 milhão de alunos matriculados em cursos do tipo, em sua maioria oferecidos por instituições com fins lucrativos. O Sinaes desenvolveu um sistema de avaliação também para esses cursos.

sificação. Algumas instituições também usaram seu desempenho para propósitos publicitários. Algumas dessas ocorrências envolveram distorções flagrantes dos dados publicados, ao ponto de o Ministério ter de intervir para pedir às instituições que removeassem os anúncios em questão.

Independentemente disso, questões mais sérias vieram à tona nos anos mais recentes, levando a preocupações com a credibilidade do sistema. O departamento de avaliação do Ministério descobriu que uma grande universidade particular estava se valendo de brechas no sistema para produzir resultados falsos: o grupo de estudantes do último ano que faria o Enade foi selecionado com o artifício de reter as notas regulares de parte dos alunos do curso até o final do semestre para o qual o Enade estava marcado. Desde então o Ministério alterou as regras para evitar tais recursos, mas não está claro se o sistema ficou livre de brechas. Houve relatos de que o Ministério teria solicitado informações de outras 30 instituições suspeitas de usar o mesmo artifício para melhorar o desempenho de seus cursos. Tudo isso foi amplamente divulgado pela mídia, já que a primeira instituição envolvida é uma das maiores do país. Os casos estão sendo investigados; entretanto, até o início de 2013, o Ministério não tinha tomado nenhuma atitude em relação às instituições envolvidas.

### CRÍTICA AO SISTEMA E PARTICIPAÇÃO DOS ESTUDANTES

Desde os primórdios do sistema federal de avaliação do ES, em 1996, quando o Ministério criou o Exame Nacional de Cursos (conhecido no Brasil como “Prova”), precursor do Enade, a participação dos estudantes tem sido um problema. No início, a União Nacional dos Estudantes (UNE) se manifestou contra todo o tipo de avaliação que envolvesse provas aplicadas aos estudantes. Muitos argumen-

Isso pode fazer o **incipiente** sistema de universidades tecnológicas seguir o clássico padrão de **desvio acadêmico**, alterando o curso de seus esforços na tentativa de se tornarem instituições “semelhantes a universidades”, algo que **afetaria as tentativas de treinar uma força de trabalho qualificada** num nível que muitos consideram ser de relevância estratégica no estágio atual do desenvolvimento econômico brasileiro.

tos foram apresentados em defesa dessa posição, alguns dos quais eram também usados por organizações de professores e especialistas em ensino, sendo eles:

- os testes não seriam capazes de medir o ensino universitário de maneira adequada;
- a participação dos estudantes equivaleria a um castigo;
- as avaliações estavam ligadas à crescente privatização do ES brasileiro, e o governo federal estaria se preparando para privatizar o sistema inteiro;
- o sistema violava a liberdade acadêmica das universidades ao impor conteúdo aos currículos dos cursos por meio da seleção de temas para o exame; e
- ainda que desenvolvessem um bom trabalho geral em termos de ensino de graduação, as instituições menos seletivas seriam prejudicadas pelo desempenho baixo de alunos com deficiências prévias de ensino.

Esses e outros argumentos certamente soarão familiares para aqueles envolvidos no atual debate internacional a respeito da avaliação dos resultados de ensino.

Quando o sistema foi reformado em 2004, muitos antigos críticos do Provão celebraram a iniciativa do Ministério, pois o Sinaes avaliava aspectos institucionais orientados por processos, e a metodologia incluía um componente de “valor agregado”, usando os resultados do Enade, na tentativa de avaliar se o curso tinha sido bem-sucedido em aumentar o conhecimento dos alunos nas áreas envolvidas no Enade. A UNE mudou sua posição oficialmente, mas muitas organizações locais e nacionais ainda apelavam para o boicote ao sistema. Do ponto de vista prático, houve casos claros de boicote por parte dos estudantes, mas há também a questão do quanto os estudantes estão envolvidos na participação no Enade. Deve ser notado que as notas individuais dos estudantes não são reveladas nem incorporadas ao seu histórico acadêmico. Um exemplo óbvio de um grande grupo de estudantes de Enfermagem que boicotou o Enade ocorreu em 2010, ocasião em que a maioria dos formandos de pelo menos quatro das maiores universidades públicas devolveu a folha de respostas em branco<sup>[7]</sup>. Além de distorcer a nota geral (já que são empregados procedimentos de normalização), os cursos dessas instituições tiveram como resultado notas

7 Foi o que ocorreu na Unicamp e nas Universidades Federais de Brasília, São Paulo e Bahia.

baixíssimas, ainda que fossem considerados entre os melhores do país pelos acadêmicos e o público. Outra evidência do boicote por parte dos estudantes: a universidade mais seletiva (Unicamp) em termos de notas do exame de admissão no Enade de 2011, apresentando a nota de admissão mais alta em 14 das 30 áreas avaliadas no exame daquele ano, não teve um único curso relacionado entre os 20 melhores (em suas respectivas áreas) no Enade realizado com formandos.

Outra questão ligada à participação dos estudantes é o fato de muitas instituições terem tentado estimular a participação dos estudantes com a oferta de prêmios àqueles que obtiverem as melhores notas (os estudantes têm acesso à própria nota), algo mais frequente entre as instituições privadas. Além disso, algumas instituições treinam seus alunos para os testes, usando versões anteriores e itens de elaboração semelhante. Ambos os casos de intervenção institucional trazem questões para os objetivos do Sinaes, em termos gerais, envolvendo a influência do sistema no ES brasileiro.

## VALIDADE DO SINAES/ENADE

O Ministério da Educação emitiu alguns relatórios técnicos explicando a metodologia empregada pelo Sinaes e o uso das notas do Enade, mas, até o momento, nenhuma análise abrangente de validade dos usos dos resultados do sistema foi desenvolvida. Além disso, foram levantados alguns pontos indicando que o sistema pode exigir certas modificações antes de podermos considerá-lo válido para a aplicação proposta: avaliar cursos e instituições para o propósito de credenciamento e regulamentação, ao menos no nível de graduação.

O primeiro ponto envolve a metodologia de definição da nota, aplicada de maneira uniforme a todas as áreas e tipos de programas, e a todos os tipos de instituição. Além de tratar da mesma maneira cursos muito diferentes, como treinamento de docentes e Engenharia, os mesmos princípios também são usados para avaliar todas as instituições, desconsiderando características acadêmicas específicas ligadas à sua missão, como a ênfase na pesquisa ou no ensino. Da mesma forma, universidades tecnológicas (que oferecem cursos vocacionais de 3 anos), outro subsistema em expansão, são tratadas no sistema de maneira semelhante às universidades, com o prevaletimento óbvio dos critérios de qualidade de tipo universitário (número de docentes com diploma de doutorado e indicadores de pesquisa, por exemplo) como referências para o restante do sistema. Isso pode fazer o incipiente sistema de universidades tecnológicas seguir o clássico padrão de desvio acadêmico, desviando seus esforços na tentativa de se tornarem instituições “semelhantes a universidades”, algo que afetaria as tentativas de treinar uma força de trabalho qualificada num nível que muitos consideram ser de relevância estratégica no estágio atual do desenvolvimento econômico brasileiro. Isso pode se aplicar a várias outras áreas dos cursos profissionais do ensino superior.

Argumenta-se também que **30 itens de avaliação não seriam suficientes para cobrir adequadamente os conhecimentos relevantes desenvolvidos num curso com diploma, já que o pequeno número de itens obrigaria os examinadores a restringir a abrangência e profundidade dos tópicos abordados.**

O MEC usa as várias notas e relatórios produzidos durante os três anos de um ciclo do Sinaes para propósitos de **credenciamento nos níveis institucional e dos cursos**. A proposta original do Sinaes (Inep, 2009) enfatizava mais a **avaliação de instituições e processos do que a nota final** ou a nota do Enade. Mas a divulgação das notas e dos resultados de credenciamento por parte da **mídia** alterou gradualmente o foco, voltando-o para tais aspectos, em especial depois que diferentes publicações impressas começaram a produzir **rankings**.

Em segundo lugar, argumenta-se que 30 itens de avaliação não seriam suficientes para cobrir adequadamente os conhecimentos relevantes desenvolvidos num curso com diploma, já que o pequeno número de itens obrigaria os examinadores a restringir a abrangência e profundidade dos tópicos abordados. Trata-se de um ponto ligado à validade do conteúdo e pode ser questionado em relação ao emprego do exame, no formato atual, para avaliar a proficiência num curso de graduação completo. Em Economia, por exemplo, debate-se entre os especialistas até que ponto o exame deve se dedicar a métodos quantitativos (econométrica), à micro e à macroeconomia, à economia desenvolvimentista e assim por diante.

Assim, semelhante à ausência de um elo entre o componente de ensino geral do Enade e os currículos dos cursos brasileiros debatida anteriormente, a questão da falta de um conteúdo claro nos exames parece ser um obstáculo para a avaliação adequada do progresso dos estudantes além dos domínios específicos de sua área de estudos. A nota de “valor agregado” desenvolvida usando as notas dos estudantes de primeiro e último ano

também merece comentários. Foi um dos pontos mais fortes avançados pelos proponentes de um novo sistema em 2004, num claro contraste com o sistema prévio existente (Provão), em favor do Sinaes. Entretanto, a maneira de computar a nota é bastante complicada, e são muitas as dúvidas em relação ao significado real da nota. Ela tem como base um modelo básico de regressão que usa as notas dos estudantes de primeiro ano, algumas outras características (como proporção de pais com diplomas do ES), para computar qual deveria ser a nota da turma de formandos. Essa previsão é então comparada ao desempenho real da turma de formandos, gerando a nota de “valor agregado”. Embora seja possível encontrar a lógica por trás desse método (Verhine et al., 2006), há muitos críticos do sistema (Schwartzman, 2008; Araújo, 2008), e não há evidências de que ele esteja de fato medindo aquilo que se propõe a avaliar.

Outro problema com o Sinaes como um todo, especialmente quando usado para propósitos de credenciamento, é o fato de o sistema não ter critério de referência para nenhuma das notas e índices computados. Assim, pode ocorrer que até

os cursos ou instituições de nota mais alta estejam apresentando desempenho abaixo dos padrões de qualidade, pois o sistema todo tem como base normas estatísticas. Por outro lado, cursos e instituições podem ter desempenho condizente com a expectativa, mas a distribuição exige que algumas sejam classificadas em níveis inferiores. Apesar das dificuldades em estabelecer pontos de corte ou faixas com critérios de referência, o Sinaes precisa usar outras metodologias disponíveis que proporcionem uma representação mais precisa do desempenho.

O sistema completo de atribuição de notas do Sinaes é bastante complicado, e pode ser questionado com justificativa em vários pontos (por exemplo, que peso é usado para as diferentes notas e por que a análise regressiva é usada dessa maneira?). Foram poucas as evidências documentadas e apresentadas ao público geral do sistema (por exemplo, estudantes, universidades, políticos etc.) para as decisões tomadas. A validade de qualquer sistema de avaliação do ensino depende muito da apresentação de argumentos convincentes para as decisões tomadas, mostrando que os resultados foram bem documentados e justificados; algo que os especialistas pedem desde a concepção do Sinaes (Schwartzman, 2008; 2012).

Outro tópico deve ser mencionado neste breve debate da validade dos sistemas de resultado de aprendizado, um tópico relacionado ao CLA: o papel desempenhado por levantamentos com base nos estudantes. Um levantamento desse tipo, que começa a ser usado por instituições de diferentes países, é o Student Experience in the Research University (SERU), desenvolvido inicialmente na Universidade da Califórnia em Berkeley, por pesquisadores do Centro para Estudos do Ensino Superior. Num relatório recente (Douglass et al., 2012), dados do SERU foram usados para desenvolver uma ava-

liação do progresso acadêmico durante os anos de faculdade. Os autores apresentam evidências de que tal levantamento pode na verdade ser muito útil para tal propósito, especialmente porque tem como base um molde censitário, tornando-o útil mesmo em se tratando de cursos específicos dentro de cada instituição. Os autores também comparam a abordagem do SERU à do CLA, oferecendo argumentos em favor do primeiro enquanto alternativa viável e interessante à abordagem de testes padronizados usada no CLA.

### **AFINAL, O SINAES/ENADE TRAZ INFORMAÇÕES ÚTEIS?**

Até o momento, o debate indica que o Sinaes/Enade não seria um sistema apropriado para avaliar o resultado de aprendizagem no ensino superior nem sua qualidade geral. Entretanto, a maioria das questões apresentadas diz respeito ao uso do sistema para a avaliação da qualidade de um único curso ou instituição por meio de uma única nota final, ou à ausência de critérios de referência para o credenciamento. Embora o Sinaes/Enade tenha várias limitações e problemas de validade (ou careça de evidências nesse sentido), para os agregados das instituições, agrupados de acordo com diferentes critérios, o sistema traz informações muito úteis. O uso dos dados relativos a todos os cursos de Engenharia, por exemplo, cujos resultados estão disponíveis a partir de 2008, é possível desenvolver um mapa detalhado do rumo seguido pelo sistema, das áreas que apresentam problemas de qualidade, quais grupos de instituições apresentam bom desempenho, se o sistema está se expandindo graças a programas mais qualificados ou não.

Como exemplo simples, tirado de um estudo desenvolvido pela Unicamp para a Petrobrás envolvendo cursos ligados à indústria de petróleo e gás, podemos ver que a maioria dos cursos mais bem qua-

**Tabela 1. Número de estudantes de último ano matriculados em cursos de engenharia, de acordo com a nota específica do Enade para esta área, 2008**

Nota do Enade	Públicas	Privadas	Total	%
5	2217	132	2349	<b>9.2</b>
4	3430	1422	4852	<b>18.9</b>
3	3525	4073	7598	<b>29.6</b>
2	2244	6565	8809	<b>34.3</b>
1	479	1564	2043	<b>8.0</b>
<b>TOTAL</b>	11895	13756	25651	

Fonte: Compilada usando dados do Enade de 2008: Inep (2008), "Resultados do Enade", Ministério da Educação, Brasil, disponível em <http://portal.inep.gov.br/planilhas-enade>.

lificados, em termos do desempenho dos estudantes do último ano, está no setor público. A **Tabela 1** apresenta o número de estudantes de último ano em cursos de Engenharia, de acordo com a distribuição da escala do Enade de 1 a 5 (sendo 5 o nível mais alto de desempenho), para instituições públicas e privadas.

A tabela mostra que mais de 90% dos estudantes em programas que obtiveram nota 5 estavam matriculados em instituições públicas, e que a grande maioria daqueles matriculados em cursos de nota 1 ou 2 (considerados "fracos" pelo critério do Ministério) estava se formando em instituições particulares. Analisando esse resultado em conjunto com a estatística segundo a qual, nos últimos 10 anos, a maior parte do crescimento nas matrículas em cursos de Engenharia ocorreu nas instituições privadas, há motivo para preocupação por parte das autoridades, dos empregadores e dos próprios estudantes.

### **O FUTURO DO SINAES E DE OUTROS SISTEMAS DE AVALIAÇÃO DOS RESULTADOS DE APRENDIZADO**

As evidências apresentadas até o momento indicam que é problemático usar o Sinaes/Enade como instrumento de regulamentação/credenciamento, levando-se em consideração os diferentes problemas envolvendo a validade dos resultados nesse contexto. Mas ainda é cedo para um veredicto final. Pode-se argumentar que seria no mínimo um instrumento válido para indicar quais áreas e setores precisam de mais controle/regulamentação por parte do governo e quais precisam de apoio e de uma avaliação mais detalhada. É preciso implementar mudanças: o sistema de notas deve usar uma metodologia com "critérios de referência"; caso contrário, o procedimento de credenciamento não faz sentido. Além disso, os diferentes indicadores não deveriam ser usados para calcular uma nota "final". Um sistema de

notas mais complexo e multidimensional seria mais válido e também mais útil.

No VI Simpósio da Associação Brasileira de Avaliação Educacional, em 2011, Schwartzman apresentou um estudo com o título “Além do Sinaes” (Schwartzman, 2011). Depois de apresentar uma ótima análise do sistema, o estudo conclui com várias recomendações, algumas das quais são apresentadas aqui:

- manter separados os diferentes indicadores, para evitar integrá-los numa nota final absolutamente obscura que traz pouca utilidade prática além de classificar cursos e instituições;
- incluir indicadores que associem os cursos aos mercados de trabalho;
- desenvolver indicadores que respeitem a diversidade das instituições e cursos;
- preparar relatórios que usem a informação de maneira que o público em geral seja capaz de compreender as questões principais sendo avaliadas e o verdadeiro significado dos indicadores; e
- remover todo o sistema da supervisão do Ministério da Educação e criar uma agência independente para lidar com os processos de avaliação e credenciamento (evitando situações de conflito de interesses, já que todo o sistema federal de instituições de ES é administrado pelo próprio Ministério).

O futuro de um sistema internacional de avaliação dos resultados de aprendizado certamente depende das descobertas do estudo de viabilidade do AHELO. Pode ser que a ideia toda seja abandonada, mas isso é improvável. É mais provável que uma variedade de sistemas

continue a coexistir, incluindo a abordagem dos levantamentos com estudantes, como exemplificado pelo SERU, que tem uma comunidade internacional de instituições de participação crescente. Essas duas metodologias básicas, a avaliação de resultados de ensino com base em testes e em levantamentos de estudantes, são na verdade complementares uma da outra. Idealmente, um curso, instituição ou sistema seria muito beneficiado pela disponibilidade de dados provenientes de ambos os sistemas, para buscarmos direções e abordagens para melhorar a qualidade do serviço. Os sistemas nacionais de avaliação das instituições de ES, como o Sinaes, devem acompanhar atentamente todos os resultados dos pilotos AHELO, usando-os para ajustar seus próprios sistemas para aliviar ou melhorar as fraquezas e lacunas.

Todos os dados produzidos pelo MEC usando o Enade são divulgados ao público. Embora condizente com as metas de transparência, isso também facilita o trabalho da mídia interessada em desenvolver “tabelas” e classificações de instituições/cursos, com todos os problemas envolvidos na manipulação ou ocultamento de informações para atender a propósitos específicos ou por falta de conhecimento.

**Algumas instituições também usaram seu desempenho para propósitos publicitários. Algumas dessas ocorrências envolveram distorções flagrantes dos dados publicados, ao ponto de o MEC ter de intervir para pedir às instituições que removesses os anúncios.**

## BIBLIOGRAFIA/REFERÊNCIAS

- AHELO** (2011), "Project update", Directorate of Education, IMHE/OECD, maio de 2011, disponível em [www.oecd.org/dataoecd/8/26/48088270.pdf](http://www.oecd.org/dataoecd/8/26/48088270.pdf).
- Arum, R. e J. Roksa** (2011), *Academically Adrift: Limited Learning on College Campuses*, University of Chicago Press, Chicago, IL.
- Balbachevsky, E.** (2004), "Graduate education: emerging challenges to a successful policy", em Colin Broock e Simon Schwartzman (orgs.), *The Challenges of Education in Brazil*, Symposium Books, Oxford, pp. 209-228.
- Balbachevsky, E. e S. Schwartzman** (2010), "The graduate foundations of Brazilian research", Higher Education Forum, Research Institute for Higher Education – Hiroshima University, Vol. 7, Hiroshima, Hiroshima University Press, pp. 85-100.
- Banta, T.** (2007), "A warning on measuring learning outcomes", *Inside Higher Education*, 26 de janeiro, disponível em [www.insidehighered.com/views/2007/01/26/banta](http://www.insidehighered.com/views/2007/01/26/banta).
- Banta, T.** (2006), "Reliving the history of large-scale assessment in higher education", *Assessment Update*, Vol. 18(4), pp. 3-4, 15.
- Brooks, D** (2012), "Testing the universities", *The International Herald Tribune*, 21-22 abril de 2012, p. 7.
- Coates, H. e S. Richardson** (2011), "An international assessment of bachelor degree graduates' learning outcomes", *Higher Education Management and Policy*, Vol. 23, No.4, OECD Publishing, Paris.
- Douglass, J., G. Thompson e C-M. Zhao** (2012), "The learning outcomes race: the value of self-reported gains in large research universities", *Higher Education Online First*, Springer, disponível em.
- INEP** (1994-2010), "Sinopses Estatísticas da Educação Superior – Graduação", Ministério da Educação, Brasil, disponível em <http://portal.inep.gov.br/superior-censosuperior-sinopse>.
- INEP** (2009), "SINAES: da Concepção à Regulamentação", Instituto Nacional de Estudos e Pesquisas Educacionais "Anísio Teixeira" (INEP), Ministério da Educação, Brasil, disponível em [www.publicacoes.inep.gov.br/detalhes.asp?pub=4389#](http://www.publicacoes.inep.gov.br/detalhes.asp?pub=4389#).
- INEP** (2008), "Resultados do ENADE", Ministério da Educação, Brasil, disponível em <http://portal.inep.gov.br/planalhas-enade>.
- Klein, S., R. Benjamin R. Shavelson e R. Bolus** (2007), "The collegiate learning assessment: Facts and fantasies", *Evaluation Review*, Vol. 31, No. 5, Sage Publications, disponível em [www.collegiatelearningassessment.org/files/CLA\\_Facts\\_and\\_Fantasies.pdf](http://www.collegiatelearningassessment.org/files/CLA_Facts_and_Fantasies.pdf).
- Nusche, D.** (2008), "Assessment of Learning Outcomes in Higher Education: A comparative review of selected practices", *OECD Education Working Papers*, No. 15, OECD Publishing, Paris, doi: 10.1787/244257272573.
- Pike, G.R.** (2006), "Value-added measures and the collegiate learning assessment", *Assessment Update*, Vol. 18(4), pp. 5-7.
- Schwartzman, S.** (2011), "Para além do SINAES", estudo apresentado no VI Simpósio da Associação Brasileira de Avaliação Educacional, Fortaleza, CE, setembro de 2011, disponível em [www.schwartzman.org.br/simon/2011\\_sinaes.pdf](http://www.schwartzman.org.br/simon/2011_sinaes.pdf).
- Schwartzman, S.** (2008), O "conceito preliminar" e as boas práticas de avaliação do ensino superior *Rev. Associação das Mantenedoras de Ensino Superior*, 38, dezembro de 2010, pp. 9-32, disponível em [www.schwartzman.org.br/simon/prelim.pdf](http://www.schwartzman.org.br/simon/prelim.pdf).
- Schwartzman, S.** (2007), "The national evaluation of courses in Brazil", Instituto de Estudos do Trabalho e Sociedade, Working Paper, disponível em [www.schwartzman.org.br/simon/provao2.pdf](http://www.schwartzman.org.br/simon/provao2.pdf).
- Souza, E. S.** (2006), "ENADE 2006: determinantes do desempenho dos alunos de Ciências Contábeis", dissertação de mestrado, Universidade de Brasília, 2008, disponível em <http://repositorio.bce.unb.br/handle/10482/3949>.
- Verhine, R., L.M.V. Dantas e J.F. Soares** (2006), "Do Provão ao ENADE: uma análise comparativa dos exames nacionais utilizados no Ensino Superior Brasileiro". *Ensaio: Aval. Pol. Públ. Educ.* V. 14, n. 52, 291-310, 2006. Versão anterior em inglês: Verhine, R. & Dantas, L.M.V. (2005), "Assessment of Higher Education in Brazil: from the Provão to Enade", documento preparado para o Banco Mundial, responsável: Alberto Rodriguez.
- US Department of Education** (2006), *A Test of Leadership: Charting the Future of US Higher Education*. Washington, DC, disponível em [www2.ed.gov/about/bdscomm/list/hiedfuture/reports/final-report.pdf](http://www2.ed.gov/about/bdscomm/list/hiedfuture/reports/final-report.pdf), acessado em 15 de maio de 2013.