

# Do Provão ao ENADE: uma análise comparativa dos exames nacionais utilizados no Ensino Superior Brasileiro\*

Robert Evan Verhine

Lys Maria Vinhaes Dantas

José Francisco Soares

## Resumo

Nesse artigo, os autores analisam, de maneira comparativa, o Exame Nacional de Avaliação do Desempenho de Estudante – ENADE, e o Exame Nacional de Cursos – ENC, mais conhecido como Provão, a partir de críticas feitas ao segundo e da capacidade do ENADE de endereçá-las. Ao fazê-lo, identificam diferenças e similaridades entre os dois exames, argumentam que ambos têm integridade técnica e apresentam contribuições para o aprimoramento do exame vigente, tais como a elaboração de uma matriz permanente e a pré-testagem dos instrumentos. Em síntese, as análises apontam para uma grande similaridade dos processos técnicos das duas abordagens, enquanto que as diferenças se apresentam mais no plano conceitual, vez que o ENADE se propõe a apresentar um indicador de diferença de desempenho e, especialmente, perde o caráter *high stakes* do exa-

me que o antecedeu. Os autores concluem que o novo exame não conseguiu, até o momento, resolver vários dos problemas identificados na abordagem anterior, mas que algumas das propostas, se implementadas, constituirão um avanço real para a utilização dos seus resultados.

**Palavras-chave:** Avaliação do ensino superior. Provão. ENADE. SINAES.

### **Robert Evan Verhine**

PhD em Educação, Universitat  
Hamburg, UH, Alemanha.  
Professor Adjunto, UFBA  
[verhine@ufba.br](mailto:verhine@ufba.br)

### **Lys Maria Vinhaes Dantas**

Doutoranda, UFBA  
[lysv@ufba.br](mailto:lysv@ufba.br)

### **José Francisco Soares**

PhD em Educação, University of Michigan  
Professor do Programa de  
Pós-Graduação em Educação, UFGM  
[francisco Soares@ufmg.br](mailto:francisco Soares@ufmg.br)

## **Abstract From the National Course Exam (Provão) to ENADE: a comparative analysis of national**

## **exams used in Brazilian High School**

*In this article, the authors comparatively analyze the National Exam for Evaluating*

\* O presente artigo é baseado em um relatório intitulado "A Avaliação da Educação Superior no Brasil: do Provão ao ENADE", financiado pelo Banco Mundial sob encomenda do seu representante Alberto Rodriguez. Os autores agradecem a colaboração de Dilvo Ristoff, Diretor de Estatísticas e Avaliação da Educação Superior no INEP, de Héligio Trindade, Presidente da CONAES, e dos membros das suas respectivas equipes que prestaram informações para esta pesquisa.

*Student Achievement – ENADE, and the National Course Exam – ENC, better known as the Provão, focusing on criticisms directed at the latter and the ENADE’s ability to address them. In so doing, they identify similarities and differences between the two exams, argue that both sustain an adequate level of technical integrity, and present suggestions for improving the present exam regarding aspects such as test specification and instrument pretesting. The analyses reveal that the technical processes adopted by the two approaches are very similar. Their differences are primarily conceptual in nature, since the ENADE, in contrast to the Provão, proposes to indicate achievement gain over time and is not high stakes in orientation. The authors conclude that the new exam has not, as yet, resolved a number of problems associated with the previous approach, but some of the current proposals, if implemented, constitute a genuine advance with respect to the utilization of the exam results.*

**Keywords:** High Education Evaluation. Provão. ENADE. SINAES.

## Resumen

### **Del “Provão” al ENADE: un análisis comparativo de los exámenes nacionales utilizados en la Enseñanza Superior Brasileña**

En este artículo, los autores analizan, de un modo comparativo, el Examen Nacional de Evaluación del Desempeño de los

Estudiantes (ENADE) y el Examen Nacional de Carreras (ENC, más conocido como Provão), a partir de críticas hechas al segundo y de la capacidad del ENADE de encaminarlas. Al hacerlo, identifican diferencias y similitudes entre los dos exámenes, argumentan que ambos tienen integridad técnica y presentan contribuciones para el perfeccionamiento del examen vigente, tales como la elaboración de una matriz permanente y el testeo previo de los instrumentos. En síntesis, los análisis señalan una gran similitud entre los procesos técnicos de los dos abordajes, siendo que las diferencias aparecen más en el ámbito conceptual, dado que el ENADE se propone presentar un indicador de diferencia de desempeño y, especialmente, pierde el carácter high stakes del examen que lo antecedió. Los autores concluyen que el nuevo examen no consiguió, hasta el momento, resolver varios de los problemas identificados en el abordaje anterior, pero que algunas de las propuestas, si fueran implementadas, constituirían un avance real para la utilización de sus resultados.

**Palabras clave:** Evaluación de la enseñanza superior – Provão – ENADES - SINAES

## Introdução

Na metade da década de 90, o governo brasileiro iniciou um processo gradual de implementação de um sistema de avaliação do ensino superior. O processo teve início em 1995 com a Lei 9.131 (BRASIL, 1995), que estabeleceu o Exame Nacional de Cursos – ENC, a ser aplicado a todos os estudantes concluintes de campos de conhecimento pré-definidos. Leis subseqüentes incluíram no sistema o Censo de Educação Superior e a Avaliação das Condições de Ensino – ACE, através de visitas de comissões externas às institui-

ções de ensino, mas o ENC, popularmente conhecido como Provão, permaneceu no centro desse sistema. Ainda que inicialmente boicotado em muitos *campi*, esse exame tornou-se parte da cultura da educação superior no Brasil. Apesar do seu crescimento (de 3 áreas de conhecimento testadas em 1995 para 26 em 2003) e da sua larga aceitação pela sociedade em geral, o Provão<sup>1</sup> foi veementemente criticado por muitos membros da comunidade acadêmica e especialistas em avaliação. Mudanças para esse exame foram amplamente debatidas durante a campanha presidencial de 2002 e, logo após o novo presidente (Luiz Inácio Lula da Silva) ter assumido o cargo, sua administração anunciou a formação de uma comissão cujo trabalho teve como objetivo sugerir alterações significativas ao sistema de avaliação vigente. Em agosto de 2003, a comissão propôs um novo sistema, chamado SINAES – Sistema Nacional de Avaliação da Educação Superior, que foi formalmente instituído através da Lei 10.861 (BRASIL, 2004), aprovada em abril de 2004. Esse novo sistema incluía uma diferente abordagem para o exame de cursos, denominado ENADE – Exame Nacional de Desempenho dos Estudantes.

O presente artigo busca descrever, comparar e analisar o modelo do ENADE em relação ao exame que se propôs a substituir, através do levantamento das principais críticas feitas ao Provão e da sua resolução, ou não, pelo ENADE. Ao fazê-lo, procura identificar as diferenças e similaridades entre o ENADE e o Provão e discute a integridade técnica dos exames.

Para cumprir tais propósitos, além da revisão da documentação existente (leis,

relatórios e artigos), foram conduzidas entrevistas com membros da equipe do INEP e da CONAES. É importante enfatizar que todos os dados levantados a respeito da implementação do ENADE referem-se à aplicação realizada em 2004. Assim, eventuais mudanças efetuadas para a aplicação em 2005 não são aqui consideradas.

O presente artigo está organizado em três seções além dessa Introdução. A primeira seção descreve e contextualiza os dois exames – o Provão e o ENADE. A segunda traz a comparação entre eles, a partir de seis das principais críticas feitas ao Provão e da tentativa do ENADE de endereçá-las. A terceira seção apresenta as considerações finais, avaliando a integridade técnica dos exames e oferecendo contribuições para a melhoria do modelo vigente.

## Os diferentes contextos do Provão e do ENADE

Conforme amplamente documentado na literatura, o Provão nasce no âmbito de globalização e neoliberalismo, em que o ensino superior é caracterizado por massificação e diversificação, por um lado, e maior autonomia institucional, por outro. Neste contexto, o estado, tanto no Brasil como no exterior, buscando conter despesas públicas e valorizando o mercado como mecanismo de alocação de recursos escassos, inicia uma política de “gerenciamento à distância”, garantindo qualidade e responsabilidade social através de processos de avaliação em lugar dos de intervenção e de controle direto.<sup>2</sup> A partir da década de 80, no cenário internacional, um modelo geral de avalia-

<sup>1</sup> Nesse artigo optou-se por utilizar a terminologia popular, Provão, em lugar de sua abreviação oficial, ENC, ao se mencionar o Exame Nacional de Cursos.

<sup>2</sup> As relações entre globalização, neo-liberalismo e a avaliação do ensino superior são discutidos no contexto internacional por Brennan; Shah (2002), Heywood (2000) e Huisman; Curry (2004) e, no contexto brasileiro, por Dias Sobrinho (2003) e Gouveia e outros (2005).

ção de ensino superior começa a emergir, caracterizado pela combinação de auto-avaliação e avaliação externa realizada por pares (BILLING, 2004). Embora este modelo não tenha se manifestado uniformemente, variando de acordo com as características específicas dos diversos contextos nacionais, é interessante observar que, dentre aproximadamente 40 países para os quais existem estudos na literatura disponível, apenas o Brasil adotou o uso de um exame nacional de cunho obrigatório.

A legislação que criou o Provão determinou que exames escritos fossem aplicados anualmente, em todo o território nacional, a estudantes concluintes de cursos de graduação. Embora precedido por outras iniciativas buscando fomentar a avaliação do ensino superior no Brasil,<sup>3</sup> foi a primeira política desta natureza a ser aplicada de forma universal e obrigatória (condicionada à liberação do diploma). Desde o início, foi planejada uma expansão gradual do número de cursos sob avaliação de modo a garantir cobertura de todas as áreas.

A política do Provão foi detalhada através do Decreto nº 2.026/96 (BRASIL, 1996), que estabeleceu medidas adicionais para a avaliação da educação superior, determinando uma análise de indicadores-chave da performance geral do Sistema Nacional de Educação Superior, por estado e por região, de acordo com a área de conhecimento e o tipo de instituição de ensino (baseada no Censo da Educação Superior). O Decreto acrescentava a avaliação institucional, cobrindo as dimensões Ensino, Pesquisa, e Extensão. Além disso, determinava que to-

dos os cursos deveriam ser avaliados através dos resultados do Provão e dos relatórios de especialistas, que verificariam *in situ* as condições de ensino em termos do currículo, da qualificação docente, das instalações físicas e da biblioteca. Desde o início, a responsabilidade pela avaliação de instituições de ensino superior (IES) e de cursos de graduação foi concedida ao INEP. O único componente do sistema que permaneceu fora do domínio do INEP foi relativo à avaliação da pós-graduação, desde 1976 sob responsabilidade da CAPES.

O aspecto *high stakes*<sup>4</sup> das avaliações conduzidas pelo INEP, que deveriam servir para orientar decisões relativas ao credenciamento institucional e ao reconhecimento e renovação do reconhecimento dos cursos, foi estabelecido oficialmente através do Decreto nº 3.860/01 (BRASIL, 2001). Apesar dessas determinações, tais aspectos nunca foram implementados. Somente em casos extremos as instituições perderam credenciamento e o processo de credenciamento periódico não chegou a ser posto em prática (SCHWARTZMAN, 2004).

Fruto da gestão de Fernando Henrique Cardoso, o Provão foi mantido pelo Partido dos Trabalhadores no seu primeiro ano de governo. Sensível ao modelo dito neoliberal atrelado à avaliação, o desagrado no novo governo tornou-se explícito quando da elaboração do Relatório Técnico do Provão 2003 (BRASIL, 2003b), do qual constam severas críticas a esse exame enquanto instrumento para medida de qualidade. Todavia, vale ressaltar que, ainda quanto à aplicação em 2003, o INEP in-

<sup>3</sup> Para informação sobre os programas de avaliação do ensino superior que antecederam o Provão, ver Brasil (2004) e Gouveia e outros (2005).

<sup>4</sup> Termo técnico da área de avaliação, sem tradução consensuada para o português, que significa "com conseqüências significativas para aquele que está sendo avaliado". O termo *low stakes* significa o contrário.

roduziu duas novidades na tentativa de endereçar essas críticas. Primeiro, publicou os resultados dos cursos em ordem alfabética, em lugar de ordem crescente de classificação nas escalas (dificultando, portanto, o *ranking* das IES). Em segundo lugar, inseriu como um dos resultados a média absoluta das áreas de conhecimento, antes relativa, baseada na curva normal.

Em paralelo à operacionalização do Provão 2003, o Ministério de Educação designou a CEA para estudar o tema da avaliação do ensino superior.<sup>5</sup> Em 27 de agosto de 2003, a Comissão divulgou a proposta preliminar para uma nova rede de avaliação, com componentes articulados e integrados, chamada Sistema Nacional de Avaliação da Educação Superior, ou mais sucintamente, SINAES.<sup>6</sup>

No final de 2003, o governo Lula aprovou a Medida Provisória 147, mudando a legislação que regia a avaliação da educação superior e, em paralelo, o MEC publicou e divulgou um documento detalhando todos os aspectos da nova abordagem de avaliação (BRASIL, 2003a). Ao fazê-lo, se ocupou em identificar várias das deficiências do exame utilizado anteriormente e em fundamentar os conceitos, princípios e características da nova proposta.

Em contraste com o sistema anterior de avaliação, a abordagem do SINAES foi pensada como verdadeiramente sistêmica e com foco na instituição. O documento acima referido, de apresentação do SINAES, obser-

vou que as avaliações institucionais deveriam fornecer análises abrangentes das dimensões, estruturas, objetivos, relações, atividades, compromissos e responsabilidade sociais, das IES e de seus cursos, nas diversas áreas de conhecimento. A proposta considerou também que os procedimentos, dados e resultados deveriam ser públicos; que a autonomia e identidade das instituições e cursos deveriam ser respeitadas, preservando-se assim a diversidade que caracteriza o setor no país; e que todos os membros da comunidade do ensino superior deveriam participar, bem como outros representantes da sociedade civil.

À Medida Provisória 147 se seguiu a Lei 10.861 (BRASIL, 2004), de 14 de abril de 2004, aprovada por grande maioria na câmara federal e no senado. Uma análise dessa lei deixa claro que nem todos os aspectos da proposta inicial para o SINAES foram aceitos e, especialmente, que várias das características da abordagem de avaliação do governo anterior foram mantidas, o que parece indicar que houve uma grande negociação entre dezembro de 2003 até a passagem da lei em abril do ano seguinte. Com ela, pode-se perceber a separação nítida entre avaliação institucional e avaliação de curso e, nesse momento, a avaliação de desempenho de estudante (ENADE) foi instituída como a terceira parte do sistema, com igual peso. Foi também instituída a Comissão Nacional de Avaliação da Educação Superior – CONAES, já prevista no delineamento do sistema proposto pela CEA, mas não contemplada na Medida Provisória.

<sup>5</sup> A Comissão Especial de Avaliação (CEA) foi formalmente instituída pelo Ministro da Educação, em 29 de abril de 2003 e apresentou sua proposta inicial para o sistema de avaliação em 27 de agosto do mesmo ano. Ela foi composta por 22 membros, incluindo 14 professores universitários, 3 estudantes, e 5 representantes do MEC (do INEP, CAPES e SESu). A CEA foi presidida pelo Prof. José Dias Sobrinho, da Universidade de Campinas, um arquiteto-chave do PAIUB, programa de avaliação anterior ao Provão, considerado por muitos como o modelo a adotar, por ter sido participativo, voluntário e centrado na auto-avaliação.

<sup>6</sup> Embora alguns autores tenham se referido ao esforço de avaliação pré-SINAES como um sistema, ele nunca foi legalmente rotulado como tal e nunca recebeu um acrônimo referenciando esse status.

Apesar das mudanças, ficaram assegurados os principais conceitos norteadores: respeito à diferença, à autonomia, à identidade, além da ênfase na missão pública e nos compromissos e responsabilidades sociais das IES. Também ficou mantido o protocolo de compromisso para aqueles que não obtivessem um conceito satisfatório e garantido o caráter regulatório do SINAES, sem excluir os aspectos avaliativo-educativos do sistema.

Além disso, enquanto o Provão se prestava a ser um mecanismo de regulação, o ENADE passaria a fornecer não só um indicador com esse fim, mas principalmente seria uma ferramenta de avaliação, através do diagnóstico de competências e habilidades adquiridas ao longo de um ciclo de 3 anos de escolarização superior, cruzado com a visão do aluno sobre sua instituição e com seu conhecimento sobre aspectos mais gerais, não relacionados a conteúdos específicos.

O estabelecimento desses princípios trouxe um grande impacto no processo de avaliação: não só os resultados deixaram de se prestar a *ranking* e a competições, como os novos objetivos do sistema o tornaram mais subjetivo, conseqüentemente, dificultando os processos de avaliação, em especial quanto ao estabelecimento de critérios e de padrões a partir dos quais decisões quanto à regulação viriam a ser tomadas.

Muitos dos aspectos de articulação entre as três peças do SINAES (avaliação institucional, de curso e do estudante) só foram melhor regulados com a publicação,

em julho de 2004, da Portaria nº 2.051 (BRASIL, 2004). Essa portaria se incumbiu de detalhar as atribuições da CONAES e do INEP (com uma grande mudança nas responsabilidades do INEP, que passa a ser apenas executor da política), bem como os procedimentos a serem usados em cada instância a ser avaliada (instituição, curso, e aluno).

Os componentes do SINAES que tratam da avaliação institucional e da avaliação de cursos não são objetos deste estudo e, assim, suas características não serão aqui detalhadas.<sup>7</sup> Quanto ao ENADE, da mesma forma que o Provão, tornou-se um componente curricular obrigatório, com registro no histórico escolar de cada estudante. Outras similaridades podem ser encontradas entre o Provão e o ENADE: os resultados individuais dos estudantes são disponíveis apenas para eles próprios; há premiação para os alunos com melhores desempenhos, por área de conhecimento; dados de perfil do aluno, do curso e da instituição, e percepção sobre a prova, são levantados em paralelo à aplicação do teste, fornecidos pelos estudantes e pelo coordenador do curso avaliado; além da já mencionada expansão gradual do sistema.

Dentre as diferenças entre o Provão e o ENADE, pode-se relacionar:

- 1 O ENADE é aplicado para estudantes ingressantes e concluintes do curso sob avaliação, desta forma incluindo nos resultados uma aproximação da noção de "valor agregado";<sup>8</sup>

<sup>7</sup> Para uma descrição de tais características, ver Brasil (2004).

<sup>8</sup> De acordo com o Diretor de Estatísticas e Avaliação da Educação Superior no INEP, o ENADE foi desenhado para medir a trajetória de aprendizagem e não para conferir desempenho (entrevista em 01/05/2005).

- 2 O ENADE avalia cada curso trienalmente, em lugar da frequência anual do Provão;
- 3 O ENADE promete ser referenciado a critério, baseando seus testes em padrões mínimos pré-determinados;
- 4 O ENADE se propõe a englobar várias dimensões em seu teste, de modo a cobrir a aprendizagem durante o curso (em lugar de apenas medir o desempenho dos alunos ao final do curso) e dar um peso maior às competências profissionais e à formação geral, com ênfase nos temas transversais;<sup>9</sup>
- 5 O ENADE visa reduzir custos através da administração dos testes a uma amostra representativa;
- 6 O ENADE divulga seus resultados de forma discreta, com pouco alarde da mídia;
- 7 O ENADE se atribui um uso diagnóstico na medida em que se diz capaz de identificar as competências não desenvolvidas pelos alunos ao longo de 3 anos de escolarização superior;<sup>10</sup> e,
- 8 O ENADE parte da premissa de que as instituições e cursos utilizarão seus resultados como ingrediente em um processo avaliativo institucional mais abrangente.

O estabelecimento de semelhanças e diferenças entre ENADE e Provão é interessante no sentido de se poder observar até que ponto o novo exame endereça as críticas feitas ao antigo, sem uma quebra de integridade técnica. Por essa razão, uma vez situados os contextos de ambas as abordagens avaliativas, a seção a seguir relaciona as principais críticas feitas ao Provão e comenta a capacidade do ENADE de solucioná-las.

## Análise comparativa entre o Provão e o ENADE a partir das críticas feitas ao primeiro

Ao comparar o Provão com o ENADE, os autores pretendem responder às perguntas: que deficiências do Provão o ENADE pretendeu suprir? Essa nova abordagem de fato trabalhou as deficiências? Para sintetizar sua resposta, optaram por relacionar a seguir seis das principais críticas feitas ao Provão e, para cada uma, comentar tecnicamente o movimento do ENADE no sentido de saná-la.

### 1ª Crítica

Faltou articulação do Provão com os demais componentes da Avaliação da Educação Superior que, por sua característica de criação, não chegou a ser implementada através de um sistema, no qual as partes contribuíssem para uma visão geral da qualidade de cada instituição.

#### Comentário

O SINAES, ainda que com algumas alterações advindas da promulgação da Lei nº 10.861/04 (BRASIL, 2004), em relação à proposta original, contempla esse problema. Cada componente do sistema contribui na composição do conceito da instância maior, ou seja, o ENADE é utilizado para a composição da nota do curso e essa é empregada para a composição do conceito da instituição.

Três ressalvas são necessárias: em primeiro lugar, os critérios de participação

<sup>9</sup> Essa diferença foi enfatizada pelos consultores do ENADE em suas apresentações na 1ª Conferência da Associação de Avaliação Educacional da América (Brasília, 11-13 jul. 2005).

<sup>10</sup> Para um aprofundamento sobre o aspecto diagnóstico do ENADE, consultar Limana e Brito (2005).

das notas do ENADE nos conceitos dos cursos e instituições ainda não foram divulgados e, a depender das decisões feitas nesse sentido, podem dissipar a importância do exame; em segundo lugar, não foram encontradas informações sobre como as informações provenientes do roteiro de auto-avaliação (ferramenta da Avaliação Institucional) serão processadas e analisadas; e por fim, em terceiro lugar, também não fica claro como outros componentes não centrais do sistema, como os conceitos da CAPES e as informações do Censo, vão contribuir para o conceito final da instituição.

## 2ª Crítica

O Provão foi imposto de fora para dentro, sem a participação das instituições a serem avaliadas. Além disso, a sociedade, de maneira geral, e a comunidade acadêmica não eram vistas como parceiros, apenas como consumidores de informação.

### Comentário

Embora o componente de auto-avaliação institucional seja a peça chave do SINAES, através da qual são respeitadas a autonomia, a identidade e a diversidade das instituições da educação superior, o ENADE, visto de maneira isolada, adota os mesmos procedimentos de representação utilizados pelo Provão. As instituições e a comunidade acadêmica são convidadas a participar ao serem representadas nas comissões de avaliação, ao terem membros fazendo parte das bancas de elaboração de testes, ao responderem ao questionário do coordenador de curso, e ao participarem de seminários e eventos com objetivo de discutir os resultados obtidos e possíveis melhoras para o exame.

Quanto à composição das comissões assessoras de avaliação, uma atenção deve ser dada ao número pequeno de representantes envolvidos, tanto para o Provão quanto para o ENADE. A indicação de membros das comissões por diversas instituições e regiões do país garante algum tipo de representatividade, mas sete participantes por área de conhecimento é um número muito reduzido (embora os autores reconheçam que comissões com uma grande quantidade de membros tendam a ser improdutivas). Essa deficiência é potencializada pela inexistência de um processo de validação externa, que envolva um número representativo de atores relevantes debruçando-se sobre as especificações produzidas por cada comissão.

Merece reflexão um segundo problema relacionado às comissões: na busca por representatividade, nem sempre o perfil técnico dos participantes é garantido e não há, nos relatórios técnicos do INEP, menção à presença de especialistas em avaliação em larga escala fazendo parte desses grupos, ainda que “experiência em avaliação” seja um item para análise do currículo dos membros das comissões. Isso pode ter um especial impacto nas especificações do ENADE por que as comissões são também responsáveis pela determinação do formato da prova.

A sociedade, de maneira geral, continua apenas sendo informada dos resultados do exame, agora com menos alarde. No entanto, é importante mencionar os planos do INEP de envolver, ainda mais, a comunidade científica no processo de discussão sobre a qualidade do Ensino Superior, através de convites formais para submissão de projetos de pesquisa baseados na utilização dos dados do ENADE. Os

projetos aprovados em termos de mérito deverão ser financiados pelo INEP. Essa estratégia tem a vantagem de gerar uma variedade bem maior de pesquisas acadêmicas e de talvez angariar um apoio junto à academia nunca desfrutado pelo Provão. Essa proposta de envolvimento, até o momento de elaboração desse artigo, ainda não tinha sido concretizada; se implementada, constitui um avanço real do ENADE em relação ao Provão.

### **3ª Crítica**

O Provão enfatizou competências finais específicas de cada área de conhecimento, deixando de lado aspectos fundamentais relacionados ao processo de ensino-aprendizagem e à formação do alunado, como compromisso profissional, ético e socialmente engajado. Além disso, o Provão se concentrou no desempenho dos alunos, em lugar de atentar para o valor agregado, gerado pelo curso no decorrer do tempo.

#### **Comentário**

Assim como o Provão, o ENADE se utiliza de quatro diferentes instrumentos para a coleta de dados: 1) os testes; 2) um questionário para levantar a percepção dos alunos sobre o teste (que faz parte do caderno de prova); 3) um questionário sobre o perfil socioeconômico-educacional do aluno (em caderno próprio, de preenchimento voluntário, e que deve ser entregue no dia da prova); e 4) um questionário a ser respondido pelo coordenador de curso sob avaliação, no qual se pedem suas impressões sobre o projeto pedagógico e as condições gerais de ensino de seu curso. No caso do ENADE, é possível comparar as respostas dos alunos ingressantes com aquelas apresentadas pelos alunos concluintes, não só em relação aos testes, mas também quanto ao perfil e à

percepção da prova, o que, contrastado com o Provão, implica o acréscimo de uma importante dimensão analítica a compor os relatórios de avaliação.

Quando comparado ao Provão, o ENADE se propõe a ampliar a gama de aspectos avaliados através do teste. O Provão manteve o foco no conhecimento profissional específico, esperado em alunos concluintes. No caso do ENADE, formação geral e componente específico representam duas dimensões distintas da prova e são analisados separadamente. Além disso, vez que o ENADE aplicou o mesmo teste a alunos ingressantes e a concluintes, as questões utilizadas deveriam cobrir os diferentes estágios (ou níveis) de aprendizagem esperados na trajetória do curso.

Considerando o novo perfil do exame, o número de questões de teste, especialmente no componente de formação geral, parece ser muito pequeno para medir o conhecimento e as competências que estão postas nas especificações. De acordo com o Resumo Técnico do ENADE (BRASIL, 2005, p. 10), os dez itens que compõem a formação geral devem medir: 1) ética profissional, 2) competência profissional, e 3) compromisso profissional com a sociedade em que vive. Ao mesmo tempo, devem avaliar a habilidade do estudante para analisar, sintetizar, criticar, deduzir, construir hipóteses, estabelecer relações, fazer comparações, detectar contradições, decidir, organizar, trabalhar em equipe, e administrar conflitos. Além dessas 12 competências, as 10 questões do teste devem considerar 11 diferentes temas, aí inclusos bio e sociodiversidade, políticas públicas, redes sociais, cidadania, e problemáticas contemporâneas.

Contudo, a análise dos itens do componente de Formação Geral oferecida nos Relatórios de Área do ENADE 2004 traz como habilidades avaliadas: analisar relações, interpretar para inferir, interpretar texto poético para associar, estabelecer pontos comuns, identificar para associar, refletir, deduzir, e efetuar leitura de gráfico. Em que medida essas oito habilidades estão relacionadas à ética, à competência e ao compromisso profissionais não fica claro, assim como as análises não aprofundam a relação do desempenho dos alunos com os temas transversais propostos. A partir da leitura dos relatórios e do Resumo Técnico (BRASIL, 2005), não é possível identificar uma lógica coerente embasando o componente de formação geral. Registra-se ainda que não houve relato de estudos que considerassem se o número de questões foi realmente adequado. Como será visto adiante, a inadequação do componente de formação geral do teste é evidenciada pelo fato de que os alunos ingressantes e concluintes atingiram, em média, escores muito próximos.

As questões do componente específico foram mais difíceis de avaliar por que diferem entre as várias áreas. Também elas foram, a priori, insuficientes para cobrir todo o escopo proposto. O problema foi exacerbado pelo fato de o número total de questões ser supostamente igual a 30.<sup>11</sup> Como consequência, o grau de cobertura do escopo proposto de cada teste deve ter variado enormemente pelas áreas.

Esses problemas também foram identificados em relação ao Provão. Entretanto, eles são mais graves no caso do ENADE porque, como já observado, esse exame

aplica o mesmo teste a ingressantes e concluintes e, por essa razão, as questões deveriam também cobrir os vários estágios de aprendizagem esperados em diferentes momentos do curso. O Provão, em contrapartida, foi menos ambicioso, vez que trazia seu foco para o conhecimento profissional específico de final de curso.

É fundamental registrar uma limitação importante identificada no processo de construção dos instrumentos, tanto para o Provão quanto para o ENADE: a ausência de pré-testagem dos itens, o que significa que um procedimento técnico fundamental para assegurar a validade e confiabilidade do instrumento não foi realizado. De acordo com fontes do INEP, essa ausência tem sido resultante de cronogramas apertados para a elaboração dos testes, muitas vezes associados à quantidade de tempo gasta no processo licitatório para contratação de empresas especializadas em avaliação, que é realizado a cada ano. Ainda que os itens e as provas, após aplicação, sejam tornados públicos e discutidos pelas comissões, agências elaboradoras, e INEP, além de serem submetidos a um crivo estatístico (parte do processo de análise dos dados), essas ações *pos-facto* não podem ser consideradas adequadas para a composição das provas devido às características de ambos os exames, especialmente em uma situação onde os itens não são reutilizados.

Como pode ser visto acima, muito da promessa do ENADE de aumento do escopo avaliado não pode ser atendida nas suas provas. Quanto ao valor agregado, o INEP o rotulou, posterior e adequadamente, como indicador de diferença de desempe-

<sup>11</sup> A análise de exemplos dos testes usados no ENADE 2004 mostrou que, ao contrário das recomendações oficiais de 30 itens, esse número variou no componente específico. Por exemplo, a prova de agronomia teve 40 itens.

no. A abordagem utilizada permite dois tipos de comparação. O primeiro é a comparação, em uma mesma aplicação, entre as médias obtidas pelo grupo de alunos ingressantes com as médias do grupo de concluintes. O segundo tipo pode ser obtido através dos resultados dos alunos ingressantes no primeiro ano do ciclo trienal de avaliação com aqueles dos alunos concluintes no terceiro ano do mesmo ciclo.

No caso do Provão, isso não era exequível, já que não havia a medição de entrada. No entanto, usando-se como medida de entrada o resultado de vestibulares seria possível produzir, em regiões restritas, uma medida do valor agregado como feito por Soares, Ribeiro e Castro (2001).

Há problemas óbvios com essas duas comparações. Quando se comparam alunos ingressantes e concluintes em um determinado ano, parte-se da premissa de que esses dois grupos são comparáveis, mas é preciso lembrar que o perfil de estudantes de um determinado curso ou instituição pode ter sido alterado ao longo do tempo. Além disso, pode haver efeito de seleção ao longo do curso (através da aprovação/reprovação dos alunos), que tende a artificialmente inflacionar o indicador de diferença de desempenho. No caso da análise dos resultados de diferentes anos, a comparação de alunos ingressantes com concluintes pode gerar incentivos indesejáveis e processos de avaliação, internos em cada curso, mais voltados para um bom indicador de diferença de desempenho. Nesse sentido, como exemplos de possíveis políticas institucionais se podem citar o encorajamento de alunos ingressantes a demonstrar um desempenho ruim no teste e/ou o aumento da seletividade no processo de

aprovação das disciplinas nos cursos avaliados, segurando no percurso aqueles alunos com desempenho que venha a “comprometer” o resultado da instituição no ENADE. Se os mesmos estudantes fossem avaliados no primeiro e no último ano do curso, esses problemas estariam sanados, mas as dificuldades logísticas associadas a um estudo em painel em larga escala são tão grandes que, embora inicialmente pensados, essa opção foi julgada não factível pela equipe do INEP. Além disso, outros problemas poderiam advir do uso do painel, como, por exemplo, o fato que os alunos, uma vez identificados, poderiam, desde o primeiro ano, sofrer pressões e políticas especiais da instituição, no sentido de garantir um bom indicador de diferença de desempenho após três anos.

Deve-se registrar ainda um outro problema com a tentativa de o ENADE calcular o valor agregado: os alunos ingressantes são avaliados ao final do seu primeiro ano letivo (com de 7% a 22% da carga horária curricular cumprida). Com isso, algum efeito institucional poderá ter sido introduzido, de modo que a diferença entre os escores dos alunos ingressantes e concluintes não capture todo o processo de aprendizagem no contexto do curso.

Problemas com comparabilidade são agravados por causa de decisão, para o ENADE 2004, de se calcular um escore final único por curso, baseado na combinação dos resultados dos alunos ingressantes e concluintes. Mais especificamente, neste cálculo, o escore padronizado dos alunos ingressantes nos itens do componente específico recebeu peso 15, o escore padronizado dos alunos concluintes nos itens do componente específico recebeu peso 60, e o es-

core padronizado de todos os alunos – ingressantes e concluintes – nos itens do componente de formação geral recebeu peso 25. Dessa maneira, os alunos concluintes contribuíram mais no cálculo dos resultados (aproximadamente 75% de concluintes contra 25% dos ingressantes, considerando-se que a proporção de ingressantes e concluintes na amostra total era aproximadamente de 2 para 1). Do mesmo modo, os componentes específicos dos testes tiveram muito mais peso que a formação geral (novamente, 75% e 25%, refletindo o número relativo de questões – 30 x 10 – que supostamente compuseram os dois blocos de itens das provas).

A razão para a composição de um resultado único, a partir dos resultados dos ingressantes e concluintes, está ligada ao fato de que a diferença entre os resultados dos dois grupos foi muito pequena, especialmente no componente de formação geral. Essa pouca diferença é surpreendente e levanta questionamentos sobre a qualidade dos itens. Simon Schwartzman (2005) chamou esse fenômeno de “O enigma do ENADE”. Uma das razões para o problema pode estar no fato, já mencionado, da utilização de um número pequeno de itens. Uma outra consideração é que os itens do componente de formação geral podem ter sido mais bem respondidos pelos alunos ingressantes por que tenderam à similaridade com as questões comuns aos vestibulares, dessa maneira favorecendo-os. Vale repetir que, sem pré-testagem, não foi possível que os desenvolvedores de teste predissessem o comportamento dos itens, fazendo adequações que se mostrassem necessárias de forma a evitar esses problemas.

Apesar dos possíveis problemas listados acima, não há dúvidas de que o foco na diferença entre os ingressantes e concluintes é o grande avanço do ENADE, quando con-

trastado com o Provão, por que essa comparação fornece algumas informações a mais sobre a aprendizagem e o perfil do aluno. O Provão, com foco exclusivo no desempenho do aluno concluinte, favorecia programas que tivessem sido mais seletivos nos seus vestibulares, ainda que esses programas não necessariamente tivessem proporcionado um ambiente mais favorável à aprendizagem. Infelizmente, esse avanço potencial do ENADE foi prejudicado pela maneira escolhida pelo INEP para o cálculo da nota final do curso. Através do uso de uma nota unificada, calculada a partir de pesos dos resultados de ingressantes e concluintes, essa comparação desaparece. Novamente, o efeito de um processo de seleção de entrada (vestibular) mais exigente, que enviesava o Provão, se apresenta no ENADE. É possível, no entanto, que esse “escore unificado” seja uma proposta temporária. Fontes do INEP mencionaram que, em três anos, uma nova abordagem será adotada, vez que, nesse momento, será possível comparar, para certos cursos, os escores dos estudantes ingressantes em 2004 com os concluintes em 2007.

## 4ª Crítica

O Provão não se mantém como indicador de qualidade, pois a não equalização dos seus instrumentos contribui para resultados instáveis e sem comparabilidade,

### Comentário

Essa crítica, presente no Resumo Técnico do ENC 2003 (BRASIL, 2003b), poderia ser aplicada inteiramente ao ENADE 2004. Não só os instrumentos não foram pensados para garantir comparabilidade, como a fórmula utilizada para a composição dos conceitos implicou resultados instáveis e, pelo menos nas áreas de Medicina e Odontologia, inflacionados, como será melhor explicado a seguir.

Em relação à atribuição dos conceitos, há duas maneiras de se transformar os escores para obtê-las. Testes referenciados a norma classificam os avaliados com base na comparação entre eles. Testes referenciados a critérios utilizam padrões pré-estabelecidos para o desempenho esperado ou desejado. Tanto o Provão quanto o ENADE argumentam que seu objetivo é determinar se os alunos demonstram ter o conhecimento e as competências que seus cursos requerem e, por essa razão, idealmente deveriam ser referenciados a critério. Em realidade, é muito difícil o estabelecimento de especificações de teste para o ensino superior, dada a complexidade das competências requeridas pelas carreiras que demandam grau superior. Um teste referenciado a critério adequado deveria conter um amplo número de questões, o que inviabilizaria sua aplicação em um período razoável. Assim, nem o Provão nem o ENADE são referenciados a critério. Esses dois exames apenas classificam os cursos avaliados a partir da comparação de seus resultados entre si, no mesmo campo de conhecimento, sem o estabelecimento de um padrão mínimo satisfatório a ser alcançado.

Tanto o ENADE quanto o Provão basearam-se nos escores de testes padronizados para dar nota e classificar os cursos avaliados<sup>12</sup>. Mas, ao fazê-lo, eles optaram por abordagens técnicas diferentes. Por exemplo, no Provão o ponto de referência para o estabelecimento da nota final de cada curso considerou os escores médios obtidos por todos os cursos de uma determinada área. No caso do ENADE 2004, o ponto de referência não considerou as mé-

dias obtidas pelos cursos em uma área, mas a distribuição dos escores brutos de todos os estudantes do campo de conhecimento sob avaliação. Como os escores dos alunos variam mais que os escores médios dos cursos, a abordagem usada pelo ENADE reduziu substancialmente o tamanho dos escores padronizados ( $z$ ), potencialmente criando distorções para a comparabilidade dos resultados.<sup>13</sup>

Um segundo problema apontado é que, no caso do ENADE 2004, a nota do curso dependeu dos melhores e piores escores dos cursos em um dado campo de conhecimento. A nota de cada curso foi determinada através da seguinte fórmula:

$$nota_i = \frac{X_i - X_{(1)}}{X_{(n)} - X_{(1)}}$$

Onde  $nota$  é a nota do curso,  $x_i$  é a nota do curso  $i$ ,  $x_{(1)}$  é a nota mínima entre todos os cursos e  $x_{(n)}$  é a maior nota entre todos os cursos.

A nota resultante foi então multiplicada por cinco e arredondada, de modo a que pudesse ser situada em uma escala de cinco níveis, variando de 01 a 05. Basear as notas nos escores mais altos e mais baixos cria distorções e instabilidades para as comparações ao longo do tempo. Uma mudança na posição relativa entre a menor nota e as outras notas das diferentes instituições fará com que o processo produza valores substancialmente diferentes. Esse fenômeno pode já ser observado em Medicina. Em 2004, o mais baixo escore de curso foi substancialmente inferior aos outros. Se esse curso

<sup>12</sup> Os autores prepararam uma nota técnica explicando os detalhes estatísticos da atribuição de notas no Provão e no ENADE, que está à disposição dos leitores interessados.

<sup>13</sup> Para calcular o escore padronizado ( $z$ ), subtrai-se o escore bruto da média de distribuição dos escores brutos e divide-se o resultado pelo desvio padrão dessa distribuição. Dessa maneira, uma maior variação aumenta o valor do denominador da equação, reduzindo, portanto, o valor do escore  $z$ .

tivesse sido excluído da análise, as notas dos demais teriam sido drasticamente alteradas para valores inferiores aos divulgados. Assim, o fato de que, em Medicina, o curso de mais baixo resultado tenha sido muito inferior aos demais resultou, no geral, que as notas da área tivessem ficado artificialmente altas. Dos 84 cursos de Medicina avaliados pelo Provão em 2003, 64 (76%) melhoraram suas notas no ENADE 2004. O percentual dos cursos de medicina classificados nas duas faixas mais altas da escala subiu de 29% (24 cursos) em 2003 para 87% (73 cursos) em 2004. O percentual de cursos nas duas categorias mais inferiores na escala caiu de 21% (18 cursos) em 2003 para zero, em 2004.

O caso de Odontologia é igualmente emblemático em relação à instabilidade dos resultados. O número de cursos alocados nas duas categorias superiores da escala subiu de 41 (32% do total) para 115 (90%), enquanto que o número de cursos das duas categorias mais baixas caiu de 36 (28%) para apenas um.

De maneira inversa ao fenômeno que aconteceu em Medicina e Odontologia, mas de acordo com o mesmo raciocínio, as notas de um determinado campo de conhecimento tornar-se-iam artificialmente baixas quando o maior escore de um dos cursos fosse substancialmente superior aos demais. Esse caso não foi identificado em nenhum dos sete campos de conhecimento avaliados em 2003 e em 2004, mas poderá acontecer no futuro, contribuindo assim para a instabilidade dos resultados ao longo do tempo.<sup>14</sup>

A instabilidade do indicador da qualidade dos cursos, proveniente tanto do Provão quanto do ENADE, é aumentada pela frequência de revisões feitas à matriz de referência de cada área de conhecimento. Não só os instrumentos em si não são equalizados ao longo dos anos, como também aquilo que é avaliado pode mudar a cada revisão. Nesse caso, é o próprio conceito de qualidade que é instável, o que dificulta a tomada de decisões das instituições e formuladores de políticas públicas no sentido de atingir um patamar, no mínimo, satisfatório.

## 5ª Crítica

O Provão seria economicamente inviável com o passar dos anos.

### Comentário

Não há um valor preciso ou estimado do custo total do Provão ou do ENADE. Em 2002, o custo relativo à contratação de agências de avaliação representou um valor de aproximadamente R\$ 100,00 por aluno (por volta de US\$ 33,00 dólares). Quanto ao mesmo custo no ENADE 2005, o INEP estima que tenha gasto R\$ 62,00 por aluno, equivalentes a US\$ 27,00. Não há informações disponíveis sobre os custos internos do Ministério da Educação com despesas com viagens para seus membros e para aqueles que fazem parte das comissões assessoras de avaliação; com seminários e eventos, com impressão de material de divulgação; com valor/hora do *staff* envolvido, com os pagamentos dos consultores responsáveis pelas análises de dados e com outros serviços.

<sup>14</sup> Os autores compararam os resultados de 2003 e 2004 em quatro campos: Medicina, Odontologia, Agronomia e Farmácia. Nesse grupo, os resultados melhoraram marcadamente nos casos de Medicina e Odontologia e moderadamente em Farmácia, permanecendo os mesmos em Agronomia. É difícil ignorar o fato de que a tendência geral de melhoria traz benefícios políticos potenciais e pergunta-se se essa consequência influenciou a decisão sobre a abordagem estatística utilizada no ENADE. Pode-se, entretanto, argumentar que não há qualquer evidência de que essas melhorias nos resultados tenham sido exploradas publicamente pelo INEP, pela imprensa ou por grupos político-partidários. Na verdade, o INEP tem sido muito "discreto" na divulgação dos resultados do ENADE.

A magnitude do esforço do ENADE 2004 foi consideravelmente menor que aquela do Provão 2003. No seu último ano de existência, o Provão foi aplicado a um total de 423.946 alunos matriculados em 5.897 diferentes cursos, distribuídos entre 26 campos de conhecimento diversos. Entre 2003 e 2004, o número de áreas avaliadas foi reduzido pela metade e o número de alunos que responderam às provas foi diminuído em dois terços. Esse fato mostra que o procedimento de amostragem cumpriu seu propósito de redução de custos, ao diminuir o número de participantes, especialmente frente às crescentes taxas de matrícula e considerando-se a inclusão de alunos ingressantes como respondentes dos testes.<sup>15</sup> Entretanto, exames em larga escala são caros por natureza e há um crescente surgimento de novos cursos e áreas, que precisarão ser contemplados no futuro. Orçamentos precisam ser garantidos no sentido de que não haja ruptura nos ciclos de avaliação ou queda da qualidade do exame, com a expansão do sistema.

## 6ª Crítica

O Provão concentrou seus esforços nos aspectos regulatórios da política, dessa maneira confundindo os conceitos de avaliação e regulação.

### Comentário

O ENADE reforça os aspectos diagnósticos do exame, contribuindo para que a dimensão Avaliação prevaleça sobre a Regulação. Entretanto, o conflito avaliação x regulação impactou o delineamento do SINAES, ressaltado, por exemplo, pelas diferenças entre a Medida Provisória 147 (BRASIL, 2003) e a Lei nº 10.861 (BRASIL, 2004), e está longe de ser resolvido. Nesse novo panorama

as referências para controles das IES por parte do governo e sociedade voltam a ser difusas. É importante reconhecer que cabe ao governo federal a supervisão e regulação da educação superior e essa função não pode ser desprezada. Talvez, com a implementação plena do SINAES, essa dimensão venha a ser contemplada com mais clareza e consistência do que até agora foi observado.

A leitura sobre seis das principais críticas feitas ao Provão e dos esforços do ENADE no sentido de resolver tais questões permite perceber que a grande maioria delas continua sem solução.

Nas Considerações Finais, a seguir, os autores apresentam algumas contribuições para o processo avaliativo.

## Considerações finais

O presente documento se propôs a descrever brevemente e a analisar o modelo do ENADE, contrastando-o ao ENC, conhecido como Provão. Após a inserção das duas experiências em seus contextos político-educacionais, uma análise técnica foi realizada a partir de seis das principais críticas feitas ao Provão e dos esforços do ENADE para saná-las.

Comparando-se os exames, foi possível observar que, quanto a questões técnicas, há muitas similaridades entre eles, especialmente em relação à elaboração e à administração dos instrumentos. Pode-se dizer que ambos adotaram procedimentos criteriosos nas etapas de construção, administração e análise dos dados e que as estratégias de divulgação respeitaram o sigilo aos resultados individuais dos estudantes, ao tempo em

<sup>15</sup> Os autores analisaram a qualidade técnica do processo de amostragem utilizado pelo ENADE –2004 e concluíram que o mesmo foi desenvolvido com rigor apropriado, considerando as limitações das informações disponíveis sobre os parâmetros relevantes.

que garantiram transparência quanto aos dados agregados (por área de conhecimento, região e dependência administrativa) e aos passos cumpridos para sua obtenção.

As principais diferenças estão intrinsecamente relacionadas à mudança de foco e objetivo do ENADE, quando comparado ao Provão. Assim, dentre elas pode-se ressaltar a inserção de um componente de Formação Geral, comum para todos os campos de conhecimento, nas provas dos estudantes; a aplicação dos testes a uma amostra de estudantes ingressantes e de concluintes; pelo envolvimento dos alunos ingressantes, a riqueza da análise de um indicador de diferença de desempenho (tanto entre ingressantes e concluintes em uma mesma aplicação, quanto entre ingressantes em um ano  $x$  e concluintes em um ano  $x+3$ ); a possibilidade da análise cruzada entre esse indicador e os diversos fatores que compõem o perfil do alunado e a percepção que tiveram da prova; a ênfase de divulgação nos aspectos avaliativos do exame, que deixa de ser *high stakes*.

Muitas das promessas do ENADE ainda não foram concretizadas. Dentre elas, algumas são pontuais, como a não-divulgação de um relatório técnico-científico do ENADE 2004, e outras são estruturais e mais sérias, como: a) as dificuldades para a análise das diferenças entre ingressantes e concluintes, criadas a partir da opção, pelo INEP, de composição de uma nota única, em lugar de uma nota dos ingressantes, e outra dos concluintes; e b) a ausência de um padrão mínimo por área. Um dos aspectos mais preocupantes quanto ao ENADE 2004, refere-se à atribuição da nota final, que implicou inflação dos cursos com conceitos mais altos na escala (mantida com cinco níveis) em algumas áreas de conhecimento, como Medicina e Odontologia.

Analisando-se os dois exames em relação à integridade técnica, pode-se afirmar que, de maneira geral, o ENADE manteve o padrão estabelecido pelo Provão. O processo de especificação dos testes adotado pelo ENADE 2004 foi similar àquele utilizado anteriormente pelo Provão. Para os dois exames, a formulação dos itens de teste, e a subsequente composição das provas foram realizadas, a partir das especificações definidas por comissões assessoras, por agências especializadas em avaliação, contratadas através de processo licitatório. O processo interno conduzido por cada agência foi descrito, com algum detalhamento, nos vários relatórios e manuais disponibilizados pelo INEP, e parece ter sido conduzido de acordo com as melhores práticas internacionais.

O grande impacto do ENADE está relacionado não a mudanças técnicas, mas principalmente, aos aspectos regulatórios do exame, que deixam de existir isoladamente. Nesse sentido, é imprescindível que sejam criados e divulgados os critérios de utilização das notas do ENADE na composição dos conceitos da Avaliação de Cursos e da Avaliação de Instituições.

Entendendo que, dos países que realizam avaliação do ensino superior, apenas o Brasil optou por uma aplicação nacional de exames escritos, por área de conhecimento, não fez parte do objetivo dessa investigação discutir a pertinência dessa escolha. Os autores partem da premissa que, adotados, os exames devem oferecer algumas contribuições para o aprimoramento do modelo atual, reconhecendo as dificuldades de delineamento de qualquer sistema de avaliação em larga escala, para efeitos diagnósticos, de regulação ou quaisquer outros necessários aos governos em geral.

Essas contribuições podem ser agrupadas em três diferentes categorias: 1) aspectos conceituais e estruturais; 2) aspectos técnicos; e 3) aspectos administrativos. Em relação aos aspectos conceituais e estruturais, uma atenção maior poderia ser dada à elaboração de matrizes de referência para cada área de conhecimento, de modo a garantir que sejam duráveis por um período mais longo de tempo que o ciclo atual. Para tanto, o processo de especificação das matrizes, por campo de conhecimento, precisa ser mais aberto, com uma ampla discussão nacional sobre seus elementos e com uma etapa de validação posterior. Também é interessante que haja mais incentivo às análises de comparação em relação às mudanças no perfil socioeconômico-educacional de alunos ingressantes e concluintes. Sugere-se incluir nesses questionários outras dimensões de pesquisa, para que sejam levantadas as contribuições da escolarização do ensino superior, não só quanto à aprendizagem dos alunos em relação aos conteúdos específicos e de formação geral, mas também em relação a atitudes e motivação, por exemplo. Por fim, defende-se uma maior articulação SINAES-CAPEs, inclusive atentando para um certo paralelismo, de modo a garantir que os dois processos avaliativos sejam complementares e não concorrentes.

Em relação a aspectos técnicos, os autores acreditam que uma grande melhoria no exame pode ser resultante da pré-testagem dos itens de prova e questionários. Essa etapa é crucial para a qualidade do instrumento e pode favorecer a representação do escopo de competências e habilidades em cada área de conhecimento, na medida em que a perda de itens por desempenho insatisfatório é menor. Além disso, é essencial para a construção de instrumentos que estes sejam comparáveis. Da mesma for-

ma, é imprescindível que a elaboração e administração de testes sejam desenvolvidas no sentido de garantir comparabilidade, em uma mesma área de conhecimento, através dos anos. Nos moldes de hoje, essa comparabilidade está comprometida.

Quanto aos aspectos administrativos, as sugestões se voltam para os processos licitatórios, que poderiam ser feitos para todo o ciclo trienal, com amarras contratuais que permitam a quebra do contrato por falta de qualidade no processo e nos produtos/ de cumprimento das cláusulas. Dessa forma, garante-se tempo para a pré-testagem dos instrumentos, a homogeneização dos materiais, e facilita-se a comparabilidade dos resultados nos anos  $x$  e  $x+3$ .

Para finalizar, é imprescindível reconhecer que nenhum processo avaliativo, especialmente considerando-se aplicação de instrumentos em larga escala, é capaz de cobrir todas as dimensões que compõem o conceito "Qualidade da Educação Superior". Acresce-se a esse reconhecimento o fato de que uma avaliação integra um ciclo de gestão que se completa com as etapas de planejamento e implementação. Por essas razões, qualquer exame ou sistema avaliativo precisa ser constantemente avaliado, para que continue a ser útil àqueles que dependem de suas informações, de caráter diagnóstico ou não. Nesse sentido, a mudança ENADE x Provão é interessante, bem como devem ser considerados os diversos aspectos mencionados acima, com o objetivo de aprimorar o exame vigente. Uma especial atenção deve ser dada à utilização dos produtos de um exame como o Provão e, depois dele, o ENADE. Todo esse esforço só se justifica se seus produtos forem efetivamente utilizados, nos diversos níveis de tomada de decisão: aluno, curso, instituição, governo e sociedade.

## Referências

- BILLING, D. International comparisons and trends in external quality assurance of higher education: commonality or diversity? *Higher Education*, Netherlands, UK, v. 47, n. 1, p. 113-137, jan. 2004.
- BRASIL. Comissão Nacional de Avaliação da Educação Superior. Portaria nº 2.051, de 9 de julho de 2004. Regulamenta os procedimentos de avaliação do Sistema Nacional de Avaliação da Educação Superior (SINAES), instituído na Lei nº 10.861, de 14 de abril de 2004. Conaes, Brasília, DF, 2004. Disponível em: <[http://portal.mec.gov.br/conaes/arquivos/pdf/portaria\\_2051.pdf](http://portal.mec.gov.br/conaes/arquivos/pdf/portaria_2051.pdf)>. Acesso em: 24 set. 2006.
- \_\_\_\_\_. Decreto nº 2.026, de 10 de outubro de 1996. Estabelece procedimentos para o processo de avaliação de cursos e instituições de ensino superior. Casa Civil, Brasília, DF, 1996. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/decreto/Antigos/D2026.htm](http://www.planalto.gov.br/ccivil_03/decreto/Antigos/D2026.htm)>. Acesso em: 24 set. 2006.
- \_\_\_\_\_. Decreto nº 3.860, de 9 de julho de 2001. Dispõe sobre a organização do ensino superior, a avaliação de cursos e instituições e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Brasília, DF, 10 jul. 2001.
- \_\_\_\_\_. Lei nº 9.131, de 24 de novembro de 1995. Altera dispositivos da Lei nº 4.024, de 20 de dezembro de 1961 e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Brasília, DF, 24 nov. 1995. Edição Extra, p. 19257. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/Leis/L9131.htm](http://www.planalto.gov.br/ccivil_03/Leis/L9131.htm)>. Acesso em: 24 set. 2006.
- \_\_\_\_\_. Lei nº 10.861, de 14 de abril de 2004. Institui o Sistema Nacional de Avaliação da Educação Superior – SINAES e dá outras providências. *Leis*, Brasília, DF, 2004. Disponível em: <<http://portal.mec.gov.br/arquivos/pdf/10861.pdf>>. Acesso em: 24 set. 2006.
- \_\_\_\_\_. Medida provisória nº 147, de 15 de dezembro de 2003. Institui o Sistema Nacional de Avaliação e Progresso do Ensino Superior. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 16 dez. 2003. Seção 1, n. 244, p.1-2.
- BRASIL. Ministério de Educação. Comissão Especial de Avaliação. *Sistema Nacional de Avaliação da Educação Superior – SINAES: bases para uma nova proposta de avaliação da educação superior*. Brasília, DF, 2003a.
- BRENNAN, J.; SHAH, T. *Managing quality in higher education*. Buckingham, UK: Open University Press, 2002.
- DIAS SOBRINHO, J. *Avaliação: políticas educacionais e reforma da educação superior*. São Paulo: Cortez, 2003.

GOUVEIA, A. B. et al. Trajetória da avaliação da Educação Superior no Brasil. *Estudos em Avaliação Educacional: revista da Fundação Carlos Chagas, São Paulo*, v. 16, n. 31, p. 101-132, jan./jun. 2005.

HEYWOOD, J. *Assessment in higher education: student learning, teaching programmes and institutions*. London: Jessical Kingsley Publishers, 2000.

HUISMAN, J.; CURRIE, J. Accountability in higher education: bridge over troubled waters. *Higher Education, Netherlands, UK*, v. 48, n. 4, p. 529-551, dez. 2004.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. *Exame Nacional de Cursos 2003: resumo técnico*. Brasília, DF, 2003b. Disponível em: <<http://www.inep.org.br>>. Acesso em: 4 dez. 2004.

\_\_\_\_\_. *Exame Nacional de Desempenho dos Estudantes – ENADE 2004: resumo técnico*. Brasília, DF, 2005.

\_\_\_\_\_. *SINAES: da concepção à regulamentação*. 2. ed. ampl. Brasília, DF, 2004.

LIMANA, A.; BRITO, M. R. F. O modelo de avaliação dinâmica e o desenvolvimento de competências: algumas considerações a respeito do ENADE. *Revista da Rede de Avaliação Institucional da Educação Superior, Campinas, SP*, v. 10, n. 2., jun. 2005.

SCHWARTZMAN, S. O contexto institucional e político da avaliação do ensino superior. In: SCHWARTZMAN, S.; DURHAM, E. R. (Org.) *Avaliação do ensino superior*. São Paulo: Ed. da USP, 1992. p. 13-27.

SCHWARTZMAN, S. O enigma do ENADE. Disponível em: <[www.schwartzman.org.br](http://www.schwartzman.org.br)>. Acesso em: 12 dez. 2004.

SCHWARTZMAN, S. The National Assessment of Courses in Brasil. In: PUBLIC POLICY FOR ACADEMIC QUALITY (PPAQ). *Research Program*. Chapel Hill: University of North Carolina, 2004. Disponível em: <[www.unc.edu/ppaq](http://www.unc.edu/ppaq)>. Acesso em: 1 fev. 2005.

SOARES, J. F.; RIBEIRO, L. M.; CASTRO, C. M. Valor agregado de instituições de ensino superior em Minas Gerais para os cursos de direito, administração e engenharia civil. *Dados, Rio de Janeiro*, v. 44, n. 2, p. 363-396, 2001.

**Recebido em:** 03/07/2006

**Aceito para publicação em:** 01/09/2006