

Avaliação da Educação Superior no Brasil – do Provão ao ENADE: um estudo meta-avaliativo

▸ Francisco José de Souza *

▸ Jurema Santos Souza **

▸ Ligia Gomes Elliot ***

▸ Lucí Hildenbrand ****

Resumo

O estudo teve como objetivo meta-avaliar o relatório apresentado ao Banco Mundial no ano de 2005, que trata das políticas públicas de avaliação da educação superior implantadas no Brasil. No procedimento meta-avaliativo foi utilizada a Lista Chave de Verificação de Scriven (2007), apoiada nos critérios estabelecidos por Davidson (2005). Como resultado, o relatório demonstrou que houve melhoria nas avaliações da educação superior, mas que alguns pontos ainda carecem de determinados avanços. Entretanto, o estudo apontou que a utilização integral de todos os padrões do *Joint Committee*, pelos autores, possibilitaria acrescentar maior qualidade aos estudos avaliativos a serem realizados no futuro.

Palavras-chave: Avaliação. Meta-avaliação. Padrões de avaliação. Lista Chave de Verificação.

* Mestre em Avaliação, Mestrado Profissional em Avaliação, Faculdade Cesgranrio. Chefe do Registro Acadêmico, Colégio de Aplicação, Universidade Federal do Rio de Janeiro; E-mail: francisco@cap.ufrj.br.

** Mestra em Avaliação, Mestrado Profissional em Avaliação, Faculdade Cesgranrio; Pós-Graduada em Sociopsicomotricidade, Faculdades Pedro II; (INES); Professora do Instituto Nacional de Educação de Surdos (INES). E-mail: juremassouza@hotmail.com.

*** PhD em Educação/Avaliação e Mestre em Estudos Latino Americanos, Universidade da Califórnia, Los Angeles (UCLA); Coordenadora do Mestrado Profissional em Avaliação do Programa de Pós-Graduação, Faculdade Cesgranrio. E-mail: ligia@cesgranrio.org.br.

**** Doutora em Ciências da Comunicação, Universidade de São Paulo – USP. Professora Adjunta do Mestrado Profissional em Avaliação, Faculdade Cesgranrio. E-mail: lucihildenbrand@yahoo.com.br.

Da avaliação à meta-avaliação

A avaliação está presente nos vários momentos da vida humana, ao deparar-se com questões como: se está frio, o que vestir? E se vai chover, é necessário levar guarda-chuva? Mesmo sem se dar conta, cada pessoa está fazendo uma avaliação, onde cada decisão, mesmo para uma resposta simples, depende de uma análise tendo em vista algum critério. Uma decisão errada pode acarretar situações desagradáveis ou constrangedoras. De acordo com Arroyo e Rocha (2010), a avaliação tem uma natureza transformadora, ao afetar a vida dos envolvidos e produzir efeito sobre eles, devendo ser utilizada como dispositivo para alavancar a tomada de decisões, tendo em vista os resultados a serem obtidos.

Partindo-se desses pressupostos, considera-se que a avaliação, para efeitos desse estudo, é um procedimento que deve ser constantemente transformado e que tem como finalidade precípua a melhoria de seus próprios processos, de suas abordagens e dos instrumentos utilizados e, principalmente, do objeto a ser avaliado. Por conseguinte, a avaliação é definida por Worthen, Sanders e Fitzpatrick (2004, p. 35), como a “identificação, esclarecimento e aplicação de critérios defensáveis para determinar o valor, a qualidade, a utilidade, a eficácia ou a importância do objeto avaliado em relação a estes critérios”, porque qualquer que seja a avaliação, ela exige escolhas e posicionamento frente a aspectos que são orientados por valores que tem como meta alcançar a qualidade.

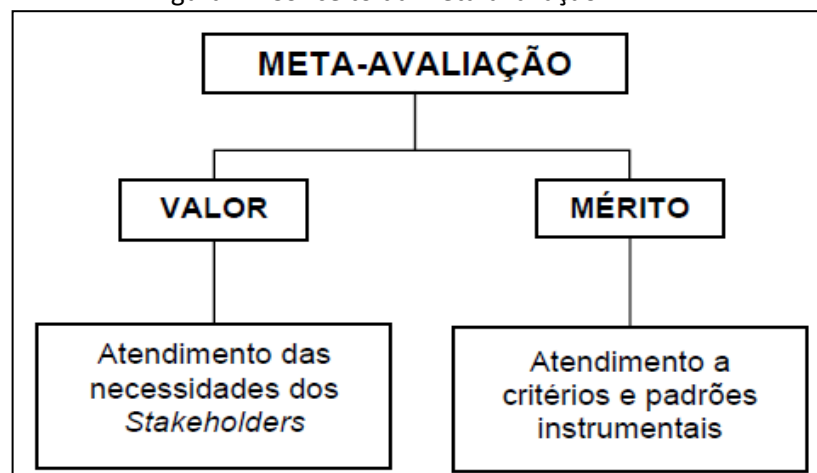
É necessário advertir, no entanto, que a avaliação realizada de forma errada pode incorrer em informações distorcidas ou falsificadas, podendo acarretar desastrosos problemas para uma organização. A importância da avaliação, na opinião de Scriven (1991 apud WORTHEN; SANDERS; FITZPATRICK, 2004), consiste no fato de que produtos e serviços mal avaliados podem custar vidas e saúde, prejudicar a qualidade de vida e desperdiçar recursos daqueles que não os têm. Sendo assim, a qualidade da avaliação deve ser motivo de preocupação, tanto para os avaliadores quanto para os clientes: para os primeiros envolve seus conhecimentos, suas convicções e sua reputação, que lhes dão a necessária credibilidade; para os demais, abrange investimentos (de tempo, político e/ou monetário) em resultados sem significância.

Logo, para evitar os transtornos de uma avaliação de má qualidade, é fundamental que ela também seja avaliada. Essa é a definição mais genérica de meta-avaliação, proposta por Scriven (1991) - “avaliação de uma avaliação” - e tem como finalidade principal o aperfeiçoamento da avaliação, podendo ser: (1) formativa, para o aprimoramento da avaliação, antes de gerar danos irreparáveis; (2) somativa, para majorar a confiabilidade dos resultados alcançados. Nesse aspecto, Davok (2006, p. 86) esclarece:

Não se presume como condição necessária a meta-avaliação ser realizada por meta-avaliadores externos, pois ela pode ser realizada pelos mesmos agentes da avaliação primária; todavia, o *Joint Committee* (1994) sugere que ela deva ser realizada por avaliadores externos, principalmente quando somativa, a fim de evitar que seja parcial ou influenciada por políticas e jogos de interesse.

A meta-avaliação deve ser coerente com os princípios de qualquer avaliação, utilizando-se de processo permanente de coleta de dados. Esses devem ser apreciados sob a luz de parâmetros para conferir valor (mensuração das necessidades de informação dos *stakeholders*) e mérito (medida em que são atendidos os critérios e padrões instrumentais estabelecidos), conforme conceito demonstrado na Figura 1.

Figura 1 - Conceito da meta-avaliação



Fonte: DAVOK (2006, p. 87).

Seus resultados devem servir de mecanismos na tomada de decisão para a melhoria da qualidade da avaliação ou para a sua *accountability* (DAVOK, 2006) e, também, auxiliar os interessados a “perceber tanto os pontos fortes como as fragilidades da avaliação realizada” (STUFFLEBEAM, 2001 apud ELLIOT, 2011, p. 944).

Corroborando a necessidade de realização da meta-avaliação, Elliot (2011, p. 960) assevera que “Em diversas instâncias, a necessidade de meta-avaliar uma avaliação, **seu relatório** ou o próprio processo avaliativo é a forma de atestar-lhes a qualidade ou de evidenciar pontos que ainda precisam ser burilados”. (grifo nosso). Vários elementos norteadores de uma meta-avaliação podem ser encontrados na bibliografia especializada. Entretanto, a referência mais conhecida talvez seja o *Joint Committee on Standards for Educational Evaluations*, que aborda padrões para a execução das avaliações e, conseqüentemente, para as meta-avaliações. Finalizando, Elliot (2011, p. 943) afirma que Stufflebeam (2001) servindo-se de sua própria definição de avaliação acrescida das dimensões descritas pelo *Joint Committee* declarou:

Meta-avaliação é o processo de delinear, obter e aplicar informação descritiva e de julgamento sobre a utilidade, a viabilidade, adequação e precisão de uma avaliação e sua natureza sistemática, competente conduta, integridade/honestidade, respeitabilidade e responsabilidade social para orientar a avaliação e divulgar publicamente seus pontos fortes e fracos.

A temática referente à meta-avaliação é muito mais ampla e até esse momento não foi esgotada. Porém, explicita suficientemente as concepções de meta-avaliação necessárias para a continuidade desse artigo. Por conseguinte, utilizando-se dos esclarecimentos e dos conceitos descritos, será apresentado o propósito desse estudo meta-avaliativo.

Objetivo e objeto do estudo

Partindo-se do entendimento dos autores citados sobre a definição de meta-avaliação somativa como método de pesquisa onde são reexaminadas etapas de avaliações conclusas, esse estudo tem como objetivo meta-avaliar o relatório da Avaliação da Educação Superior no Brasil: do Provão ao ENADE¹, onde serão tratadas as abordagens relevantes ao estudo meta-avaliativo, conforme salientado por Elliot (2011): Lista Chave de Verificação da Avaliação de Scriven (2007), que serve para nortear pontos de verificação do objeto avaliado, e Padrões de Avaliação do Comitê Conjunto (JOINT COMMITTEE..., 2011).

¹ ENADE se refere ao Exame Nacional de Desempenho do Estudante, destinado a universitários, no país.

O objeto a ser meta-avaliado é um relatório que se refere à avaliação da educação superior no Brasil: documento apresentado ao Banco Mundial, datado de dezembro de 2005, foi elaborado por Robert Evan Verhine², Doutor em Educação pela Universitat Hamburg, é também membro do Conselho Diretor da Comissão Fulbright do Brasil, da Comissão Editorial do CNPq, do Conselho de Ciência e Tecnologia do Estado da Bahia e membro-fundador da Academia de Ciências da Bahia, além de ter atuado como Pró-Reitor de Ensino de Pós-Graduação da Universidade Federal da Bahia (UFBA), foi Presidente da Comissão Nacional de Avaliação do Ensino Superior e Vice Presidente da Associação Brasileira de Avaliação Educacional; e por Lys Maria Vinhaes Dantas³, Doutora em Educação pela Faculdade de Educação da UFBA, professora adjunta da Universidade Federal do Recôncavo da Bahia, foi assessora da Diretoria Científica da Fundação de Amparo à Pesquisa do Estado da Bahia e coordenadora do Projeto de Avaliação Externa da Educação Pública Fundamental na Bahia, possuindo experiência nas áreas de Administração e Educação, com ênfase em Avaliação de Sistemas, Planos e Programas Educacionais, atuando principalmente nos temas da avaliação educacional, da política pública educacional, das políticas públicas de educação superior e da gestão de pesquisa científica.

O documento é composto de 48 páginas e traz em seu bojo o histórico dos modelos de avaliação utilizados no país no fim do século passado até os dias atuais, organizado em seis tópicos: (i) o Provão⁴ e os demais componentes da avaliação do Ensino Superior, de 1995-2003; (ii) Sistema Nacional de Avaliação do Ensino Superior (SINAES), de 2004 até o presente), incluindo-se nesse caso a proposta, seus conceitos e princípios, a legislação, os componentes, a avaliação institucional, a avaliação do curso, o ENADE, a Governança, a Comissão Nacional de Avaliação da Educação Superior (CONAES) e o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP); (iii) análise técnica comparativa entre o Provão e o ENADE, contendo a elaboração de instrumentos, a administração de instrumentos, análise

²: Currículo Lattes Disponível em:

<<http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4787528D2>>.

Acesso em: 19 abr. 2017.

³ Currículo Lattes Disponível em:

<<http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4786160J7>>

Acesso em 19 abr. 2017.

⁴ O Exame Nacional de Cursos (ENC) foi chamado de Provão.

de dados, a comunicação e divulgação dos resultados e os custos operacionais; (iv) considerações finais; e (v) siglas utilizadas nesse documento. O relatório pode ser acessado no seguinte site disponível em:

<<http://www.isp.ufba.br/avalia%C3%A7%C3%A3o%20da%20Ed%20Superior%20do%20Provao%20ao%20ENADE.pdf>>.

Apesar de não ser o foco desse estudo, é interessante saber que o documento originou um artigo denominado “Do Provão ao ENADE: uma análise comparativa dos exames nacionais utilizados no Ensino Superior Brasileiro”, tendo como autores os mesmos que elaboraram o relatório e também José Francisco Soares. O artigo se encontra disponível em: <<http://www.scielo.br/pdf/ensaio/v14n52/a02v1452.pdf>>.

Desenvolvimento do estudo

Em qualquer avaliação, e, por conseguinte meta-avaliação, os diversos interessados, ou *stakeholders*, devem ser identificados, desde os responsáveis pela tomada de decisões até os que apenas desejam obter informação sobre o objeto avaliado, porque, considerando a afirmação de Patton (1986 apud PENNA FIRME; LETICHEVSKY, 2010) de que o valor da avaliação é expresso na utilização de seus resultados, quanto maior for o comprometimento desses atores, tanto melhor será essa utilização. Penna Firme e Letichevsky (2010) apontam ainda que essa concepção fomenta a imprescindível necessidade de capacitar educadores, líderes, dirigentes e profissionais de diversas áreas para melhor utilização da avaliação nos diversos contextos disciplinares. Assim, além dos *stakeholders*, esses são os principais interessados no desenvolvimento desse estudo, que busca trazer à tona métodos que visem à melhoria da qualidade das avaliações e, como consequência, o melhor aproveitamento dos resultados alcançados.

Implementado pelos autores nos meses de abril e maio de 2017, na sede da Faculdade Cesgranrio, município do Rio de Janeiro, RJ, o estudo meta-avaliativo teve início, primeiramente, com a escolha do relatório a ser meta-avaliado. Por intermédio de pesquisas realizadas nas diversas bases de dados existentes (SciELO, IBICT), a escolha recaiu sobre aquele que é o objeto dessa meta-avaliação.

Após, foi conduzido um levantamento bibliográfico, nas mesmas bases citadas, acerca do tema em questão, buscando-se artigos (BALBIM et al., 2012; FURTADO; LAPERRIÈRE, 2012; HEDLER; TORRES, 2009), dissertação (ADLER, 2002) e teses (DAVOK,

2006; PINTO, 2015) de diferentes autores; Por fim, considerando que a meta-avaliação é uma avaliação e que requer a utilização dos mesmos conceitos, foi definido o instrumento de coleta de informação a ser utilizado. Nesse aspecto, Elliot (2012, p. 13) esclarece:

Um instrumento de pesquisa ou avaliação consiste no recurso usado para coletar a informação de interesse sobre uma variável, característica, categoria ou dimensão do objeto, ou ainda evidências de indicadores. São os dados coletados pelo instrumento que auxiliam o avaliador ou pesquisador a acompanhar o desenvolvimento do objeto ou fenômeno focalizado, a obter informações mais precisas sobre ele, e a tirar conclusões sobre determinadas características suas.

Decidiu-se, então, após a escolha do relatório, pela utilização de uma adaptação da Lista Chave de Verificação da Avaliação, proposta por Scriven (2007 apud ELLIOT, 2011), para a realização do estudo meta-avaliativo. Cabe esclarecer que a adaptação se dá pela supressão da coluna “Justificativa do Julgamento”, que será apresentada juntamente com o relato da aplicação do instrumento para cada evidência encontrada. Para isso, foram utilizados os padrões e os critérios apresentados no Quadro 1, buscando-se determinar quais aqueles percebidos no relatório.

Quadro 1 – Adaptação da Lista Chave de Verificação da Avaliação

Pontos de Verificação		Julgamentos (A – E)
Preliminares	I. Sumário Executivo	
	II. Prefácio	
	III. Metodologia	
Fundamento	1. Antecedentes e Contexto	
	2. Descrições e Definições	
	3. Consumidores	
	4. Recursos	
	5. Valores	
Subavaliações	6. Processo	
	7. Resultados	
	8. Custos	
	9. Comparações	
	10. Exportabilidade	
Conclusões	11. Síntese	
	12. Recomendações e Explicações	
	13. Responsabilidade e Justificativa	
	14. Relatório e Apoio	
	15. Meta-avaliação	

Fonte: SCRIVEN (2007 apud ELLIOT, 2011).

A legenda, enunciando os critérios utilizados nos julgamentos, foi elaborada por Davidson (2005 apud ELLIOT, 2011) e complementa a Lista de Verificação:

- (A) atinge todos os aspectos principais deste ponto e os expressam de forma clara e concisa;
- (B) atinge a maioria dos aspectos relativos a este ponto, mas não cobre um ou dois deles, ou trata dos ingredientes principais, mas não são 100% claros;
- (C) trata do ponto, mas perde alguns aspectos cruciais, ou perde vários aspectos importantes, ou é um tanto desorganizado ou pouco claro;
- (D) tem um ou dois elementos que parecem tratar implicitamente do ponto, mas o tratamento é deficiente;
- (E) perde o ponto totalmente.

Definidos os padrões de avaliação a serem utilizados, serão apresentados na sequência os resultados da meta-avaliação realizada. Antes, porém, de avançar sobre a apresentação dos resultados, compete informar que nesse estudo não foi considerado o Ponto de Verificação número 10 - Exportabilidade, que faz parte do nível Subavaliações, posto que se refere, segundo Scriven (2007), “à capacidade de os componentes da avaliação terem valia para uma futura utilização em outro local.” (ELLIOT, 2011). No caso do relatório meta-avaliado, os componentes são muito específicos, e não se aplicam a outro caso de estudo.

Resultados

No primeiro nível, ou componente da Lista de Verificação de Scriven, denominado Preliminares, o primeiro Ponto de Verificação, (I) Sumário Executivo, deve fornecer, de forma sintética, uma visão geral dos resultados relativos aos últimos cinco pontos da Lista Chave de Verificação da Avaliação, que correspondem àqueles numerados entre 11 e 15.

Dessa forma, nas considerações finais são apresentados, de forma resumida, os principais pontos apontados pelos autores no relatório meta-avaliado, trazendo o objetivo do documento de descrever e avaliar o modelo do ENADE, contrastando-o ao Exame Nacional de Cursos (ENC), conhecido como Provão. Para tanto, tais modelos foram inseridos em seus contextos político-educacionais, realizando-se análise técnica a partir de cinco dimensões: (i) a elaboração do instrumento; (ii) sua aplicação; (iii) os processos

e critérios utilizados na análise de dados; (iv) as estratégias de comunicação e de divulgação dos resultados e (v) os custos operacionais (VERHINE; DANTAS, 2005).

Complementando, é realizada na continuidade a apresentação sintética das respostas dadas pelos autores do relatório a cada uma das quatro perguntas que o nortearam:

“(i) Quais as diferenças e similaridades entre o ENADE e o Provão?” (VERHINE; DANTAS, 2005, p. 4).

Nas questões técnicas, foram verificadas muitas similaridades entre as duas formas de avaliação, principalmente na construção e administração dos instrumentos avaliativos, utilizando-se de procedimentos criteriosos criação, administração e análise dos dados e as táticas de divulgação, respeitando-se a confidencialidade dos resultados alcançados pelos estudantes e garantindo-se a lisura dos dados e das etapas para o seu atingimento. A diferença principal está na mudança de foco e do objetivo do ENADE, comparado ao Provão, com a inclusão de componente de Formação Geral, comum a todas as áreas de conhecimento nas provas, na aplicação dos testes a uma amostra de estudantes que ingressam e que concluem o curso e pelo comprometimento desses mesmo alunos, dentre outras. (VERHINE; DANTAS, 2005).

“(ii) O ENADE manteve a integridade técnica do Provão?” (VERHINE; DANTAS, 2005, p. 4).

Manteve, principalmente quanto aos fatores regulatórios, que não estão isolados, sendo necessário, no entanto, a elaboração e difusão de critérios de utilização das notas no ENADE na constituição dos conceitos da Avaliação de Cursos e das Instituições. (VERHINE; DANTAS, 2005).

“(iii) Que deficiências do Provão o ENADE pretendeu suprir? Essa nova abordagem de fato trabalhou as deficiências?” (VERHINE; DANTAS, 2005, p. 4).

Sinteticamente serão apresentadas críticas ao Provão e à dinâmica do ENADE para saná-las, ou não:

(a) faltou conexão do Provão com os demais componentes da avaliação da Educação Superior. Diferentemente, o ENADE é usado para compor a nota do curso, que por sua vez, é utilizada na composição do conceito da instituição;

(b) o Provão foi imposto, sem a participação das instituições avaliadas, tendo a sociedade e a comunidade acadêmica apenas como consumidores de informação. O ENADE, isoladamente, apresenta o mesmo comportamento;

(c) o Provão visou às competências finais, desprezando enfoques do processo ensino-aprendizado da formação discente no decorrer do curso. Contrariamente, o ENADE introduziu elemento de formação geral nos testes, com a responsabilidade de atingir todo o caminho do estudante;

(d) os resultados do Provão se devem ao fato de não haver um nivelamento dos seus instrumentos, gerando resultados inflacionados, instáveis e sem similaridades, não se considerando como indicador de qualidade. Essa crítica, que está no Relatório Técnico do ENC 2003, pode ser empregada, também, ao ENADE, porque não apenas os instrumentos não foram gerados para se ter comparabilidade como a maneira de estruturar os conceitos inflacionou cursos nos níveis superiores da escala de resultados, como Medicina e Odontologia;

(e) tendo em vista o modelo censitário utilizado pelo Provão, que o tornaria inviável financeiramente no futuro, o ENADE se mostrou menos custoso servindo-se de procedimentos amostrais;

(f) os conceitos de avaliação e regulação foram confundidos pelo Provão ao serem convergidos para questões regulatórias da política. Por sua vez, o ENADE contribui para que a avaliação predomine sobre a regulação, reforçando os aspectos de diagnose. Porém esse conflito avaliação x regulação gerou impacto no escopo do SINAES, pois as relações de controle das instituições de ensino superior pelo governo e pela sociedade voltam a ser obscuras, sendo importante reconhecer o papel de supervisão e de regulação superior que não pode ser desprezada. (VERHINE; DANTAS, 2005).

“(iv) Como o ENADE pode ser melhorado?” (VERHINE; DANTAS, 2005, p. 4).

São recomendadas algumas práticas em relação às perspectivas estruturais e conceituais, como:

Garantia de condições de trabalho para que a CONAES funcione de maneira independente, mas articulada com os prazos e cronogramas do INEP, de modo a não prejudicar a execução da avaliação. Maior articulação governo federal x estados, de modo a promover e sensibilizar a participação das instituições de educação superior estaduais nas aplicações do exame. Maior articulação SINAES - CAPES,

de modo a garantir que os dois processos avaliativos sejam complementares e não concorrentes. Elaboração de matrizes de referência para cada área de conhecimento que sejam duráveis por um período mais longo de tempo que o ciclo atual. (VERHINE; DANTAS, 2005, p. 44).

Todavia, o processo de especificação das matrizes, por campo de conhecimento, precisa ser mais aberto, com uma discussão nacional sobre seus itens, e com uma etapa de validação. Incentivo às análises de comparação em relação às mudanças no perfil socioeconômico-educacional de alunos ingressantes e concluintes.

Ao responder as perguntas norteadoras da avaliação Provão X ENADE percebeu-se que nenhum dos processos avaliativos está habilitado a atingir todas as variáveis que compõem a concepção da Qualidade da Educação Superior, pois a avaliação pertence a um ciclo, que tem como percurso o planejamento e a implementação. Dessa forma, qualquer modelo de avaliação deve ser constantemente avaliado para ser considerado útil àqueles que se utilizam das suas informações. Dessa forma, pode-se considerar interessantes as alterações ENADE x Provão, e devem ser levados em conta os diversos enfoques já mencionados, buscando-se o aprimoramento do exame. Particular atenção deve ser dada ao aproveitamento dos produtos do Provão e, posteriormente, do ENADE, pois todo o esforço depreendido só é justificável se esses produtos forem, de fato, usados nos diversos níveis decisórios: aluno, curso, instituição, governo e sociedade. (VERHINE; DANTAS, 2005).

Ainda no primeiro nível da Lista de Verificação, o segundo Ponto de Verificação, (II) Prefácio, comunica “quem encomendou a avaliação e as razões principais, quais as questões avaliativas e quais os interessados ou públicos que aguardam os resultados da avaliação e/ou deles necessitam.” (ELLIOT, 2011, p. 945). Nesse ponto, são mostrados pelos autores o objetivo de descrever, comparar e avaliar o ENADE em relação ao Provão e as questões avaliativas, chamadas de questões norteadoras pelos autores do relatório:

1 - Quais as diferenças e similaridades entre o ENADE e o Provão? 2 - O ENADE manteve a integridade técnica do Provão? 3 - Que deficiências do Provão o ENADE pretendeu suprir? Essa nova abordagem de fato trabalhou as deficiências? 4 - Como o ENADE pode ser melhorado? (VERHINE; DANTAS, 2005, p. 4).

Quanto aos interessados e quem encomendou a avaliação, não estão explicitados no texto. Porém, pode-se inferir pelo título que a encomenda foi feita pelo Banco Mundial, sendo seus acionistas, os interessados. Além desses, têm-se também como igualmente interessados a sociedade, as instituições de ensino avaliadas, seus alunos e professores, o governo, dentre tantos outros.

O terceiro e último Ponto de Verificação do nível das Preliminares, (III) Metodologia, apresenta os procedimentos utilizados (incluindo o desenho e a abordagem adotados pela avaliação, entre outros procedimentos). Busca-se o entendimento do que é necessário para progredir nos procedimentos metodológicos da avaliação e como foram desenvolvidas, ou o que era preciso e não foi feito no processo avaliativo (ELLIOT, 2011).

Diante disso, nas etapas metodológicas apontadas, segundo os autores do relatório, buscou-se descrever, comparar e avaliar o modelo do ENADE em relação ao exame que se propôs a substituir, levando-se em conta as principais etapas das duas experiências avaliativas, dentre as quais o desenho conceitual, governança, amostragem, elaboração e administração do teste, análise de dados e utilização dos resultados, buscando responder as perguntas norteadoras citadas. Para realizar a análise, fez-se um levantamento preliminar de informações com base na documentação existente (leis, relatórios e artigos), disponível em livros e revistas especializadas e no site do INEP, caracterizando-se, nesse caso, uma abordagem documental.

A seguir, um quadro comparativo Provão x ENADE foi composto e, a partir dessa primeira etapa, foi elaborado um roteiro para as entrevistas que foram conduzidas com membros da equipe do INEP e da CONAES, para esclarecimento de pontos não apresentados pela documentação disponível. Outras informações e aprofundamentos foram buscados na participação em eventos, como a 1ª Conferência da Associação de Avaliação Educacional das Américas e o lançamento do SINAES na UFBA, e do acompanhamento de discussões acadêmicas.

Na terceira fase do trabalho, após a sistematização dos dados coletados (já enriquecidos pela divulgação do Resumo Técnico do ENADE 2004), foi convidado um psicometrista (José Francisco Soares) para elaborar um parecer sobre as decisões do ENADE 2004 quanto aos processos de análise dos resultados e sua alocação nas escalas.

A última etapa do trabalho foi a elaboração do presente artigo, que está organizado em quatro seções. A primeira seção descreve o Provão e o contextualiza no sistema de avaliação do qual fez parte. A segunda apresenta a descrição do modelo SINAES-ENADE, indicando suas características-chave e modos de operação. Traz, na terceira seção, a comparação ENADE-Provão, dividida em cinco subseções: “(a) Elaboração dos instrumentos; (b) Administração de instrumentos; (c) Análise de dados; (d) Comunicação e divulgação dos resultados; (e) Custos operacionais” (VERHINE; DANTAS, 2005, p. 4).

Por fim, na quarta e última parte, Considerações Finais, o documento apresenta a avaliação sobre a integridade técnica do ENADE, oferecendo sugestões para sua melhoria.

A partir dessa fase, onde acontece de fato a meta-avaliação, são exibidos para cada nível os seus respectivos Pontos de Verificação da Lista de Scriven (2007), o julgamento atribuído a cada ponto a partir dos critérios sugeridos por Davidson (2005) e o fragmento do relatório atestando o julgamento efetuado.

Assim, no segundo nível, Fundamentos, numerados de 1 até 5, são tratados alguns atributos da avaliação. O primeiro, (1) Antecedentes e Contexto, no qual é considerada a razão de existir do objeto, podendo ser um projeto, um programa, material, entre outros (ELLIOT, 2011), foi julgado satisfatório e atribuído o conceito máximo **A**. O texto que fundamenta a avaliação pode ser encontrado na página 5 do relatório:

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 9.1311 que estabeleceu o Exame Nacional de Cursos (ENC), a ser aplicado a todos os estudantes concluintes de campos de conhecimento pré-definidos. (VERHINE; DANTAS, 2005, p. 3).

O segundo Ponto de Verificação, (2) Descrição/definição do objeto, também foi julgado procedente, recebendo conceito **A**, de melhor julgamento, e pode ser localizado na página 5:

O presente artigo busca descrever, comparar e avaliar o modelo do ENADE em relação ao exame que se propôs a substituir. A análise procurará levar em conta as principais etapas das duas experiências avaliativas, dentre as quais o desenho conceitual, governança, amostragem, elaboração e administração do teste, análise de dados e utilização dos resultados [...] O Provão foi criado em 1995 através de uma lei federal (Lei nº 9.1319) e detalhado, um ano mais tarde, através

do Decreto nº 2.02610. Essa legislação determinou que exames escritos deveriam ser aplicados anualmente, em todo o território nacional, a estudantes concluintes de cursos de graduação. Foi planejada uma expansão gradual do número de cursos sob avaliação de modo a garantir cobertura de todas as áreas. (VERHINE; DANTAS, 2005, p. 3-5).

Da mesma forma que o anterior, o terceiro Ponto de Verificação, (3) Consumidores, que são os principais interessados, atinge todos os aspectos apreciados e também foi considerado **A**, como mostrado na página 5:

Os resultados agregados de cada curso avaliado deveriam ser encaminhados para as instituições que os tinham oferecido, mas nenhum dado sobre os resultados individuais dos alunos deveria ser fornecido. A participação dos estudantes no exame era obrigatória e aqueles que se recusassem a fazê-lo não obteriam o diploma de graduação.” (VERHINE; DANTAS, 2005, p. 5).

O conceito atribuído ao Ponto de Verificação (4) Recursos, que se caracteriza como aquele disponível para manter o objeto e ajudá-lo a ser eficaz em seus propósitos, foi **D**. A definição desse conceito é: “Tem um ou dois elementos que parecem tratar implicitamente do ponto, mas o tratamento é deficiente.” (DAVIDSON, 2005 apud ELLIOT, 2011, p. 945). Os autores não mencionam diretamente os recursos (humanos, monetários etc.) utilizados, fazendo apenas uma menção a especialistas e ao final, nos resultados, alguns custos financeiros.

O último Ponto de Verificação desse nível, (5) Valores, está relacionado diretamente ao padrão adotado pela avaliação: valores relativos a requisitos legais, éticos, profissionais, culturais, históricos, tradicionais, científicos, tecnológicos, políticos, econômicos, o que não esgota a lista de possibilidades (SCRIVEN, 2007 apud ELLIOT, 2011). Os valores para a Educação aparecem implícitos na avaliação, porém, na redação do estudo, os autores expõem a importância do objeto estudado, mostrando os embasamentos legais, o ambiente político, e a importância de todo processo de aperfeiçoamento na aplicação das avaliações a serem realizadas pelo Governo. Assim, recebeu no julgamento o conceito **A**, que indica o atingimento pleno do critério mais completo de Davidson (2005).

O terceiro nível, Subavaliação é onde, segundo Elliot (2011),

São conduzidas ao processo e aos resultados ou impactos da avaliação, aos custos (considerados pelo autor como elemento-chave para determinar valor), às comparações de um programa com outro mais efetivo e, também, à capacidade de os componentes da avaliação terem valia para uma futura utilização em outro local. A relevância geral da avaliação também faz parte desse nível, incluindo o que está indo bem, o que está faltando, qual o melhor uso dos recursos, respeitando-se a relação custo-efetividade.

Nesse nível, os Pontos de Verificação estão numerados de seis até 10, onde o primeiro, (6) Processo, não foi demonstrado totalmente, recebendo com isso o conceito **B**, ou seja, não atingiu um ou dois aspectos considerados importantes. No entanto, o relatório trata de alguns itens pertinentes:

Levantamento preliminar de informação com base em documentação existente (leis, relatórios, artigos etc.). Após a sistematização dos dados coletados (enriquecido pela divulgação do Resumo Técnico do ENADE 2004), um psicometrista foi convidado a elaborar um parecer [...] (VERHINE; DANTAS, 2005, p. 4).

O Ponto de Verificação (7) Resultados ou Impactos da avaliação foi julgado coerente e recebeu o conceito **A**, pois traz no texto:

O ENADE parte da premissa de que as instituições e cursos utilizarão seus resultados como ingrediente em um processo avaliativo institucional mais abrangente. [...] Tanto o Provão quanto o ENADE ofereceram prêmios aos estudantes com melhor desempenho. Em 2004, o INEP ofereceu bolsas CAPES para pós-graduação ao estudante com o escore mais alto em cada uma das 13 áreas avaliadas. Uma política similar já tinha sido posta em prática pelo Provão desde 2000. (VERHINE; DANTAS, 2005, p. 23-38).

No Ponto de Verificação (8) Custos, também recebeu o conceito **A**, pois sendo considerado elemento-chave na determinação de valor, os autores apontaram de forma pertinente:

O ENADE parece ser muito mais econômico que o Provão: custo per capita mais baixo e testou um número menor de estudantes. Não há informações sobre os custos internos do MEC com despesas de viagens [...]. 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 outros.” (VERHINE; DANTAS, 2005, p. 40).

Da mesma forma que o anterior, o Ponto de Verificação (9) Comparações recebeu o julgamento **A**, dado que se refere à comparação entre um programa e outro mais efetivo e é apontada adequadamente pelos autores na página 41 do texto:

Quais as diferenças e similaridades entre o ENADE e o Provão? 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*. (VERHINE; DANTAS, 2005, p. 41).

Como já fora citado, o Ponto de Verificação (10) Exportabilidade não se aplica a esse estudo.

O quarto e último nível da Lista de Verificação de Scriven (2007), Conclusões, está numerado de 11 até 15, e diz respeito às recomendações e explicações, que são opcionais, e a responsabilidade dos resultados e justificativas, mas nem sempre as avaliações abordam informações dessas características. Os relatórios trazem a divulgação dos resultados, os formatos e os destinatários. Enfim, o último Ponto de Verificação trata, especificamente, da meta-avaliação, cuidando dos pontos fortes e fracos da avaliação e as conclusões acerca de padrões de avaliação: utilidade, viabilidade, adequação e precisão (ELLIOT, 2011).

O primeiro Ponto de Verificação desse nível, (11) Síntese, foi julgado totalmente procedente e recebeu conceito **A**, conforme demonstrado no relatório:

Após a inserção das experiências em seus contextos político-educacionais, realizou-se análise técnica a partir de cinco dimensões: elaboração do instrumento; aplicação; processos e critérios utilizados na análise de dados; estratégias de comunicação e divulgação dos resultados; e custos operacionais. Para finalizar, essa seção apresenta a síntese das respostas a cada uma das quatro perguntas que nortearam o trabalho. (VERHINE; DANTAS, 2005, p. 41).

O relatório apresenta recomendações na página 44:

IV. Como o ENADE pode ser melhorado? (a) - Em relação a aspectos estruturais e conceituais, recomenda-se: Garantia de condições de trabalho para que a CONAES funcione de maneira independente (infraestrutura física e de pessoal, possibilidades de reuniões mais frequentes, contratação de algumas pessoas de apoio que sejam permanentes e que façam as articulações entre os diversos membros), mas articulada com os prazos e cronogramas do INEP, de modo a não prejudicar a execução da avaliação. (VERHINE; DANTAS, 2005, p. 44).

Além das recomendações, são apresentados críticas e comentários referentes às perguntas norteadoras: “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 vistos como parceiros, apenas como consumidores de informação.” (VERHINE; DANTAS, 2005, p. 41). Dessa forma, o Ponto de Verificação (12) – Recomendações e Explicações, por sua abrangência e pertinência, recebeu conceito **A**.

O Ponto de Verificação (13) Responsabilidade e Justificativa não foram encontrados no texto. Entretanto, o ENADE, bem como o seu antecessor – Provão, é uma política de Estado. Fica implícita a sua responsabilidade, sem, no entanto, cumprir totalmente o que o ponto sugere, sendo-lhe atribuído, portanto, conceito **C**.

O Ponto de Verificação (14) Relatório e Apoio foi atingido integralmente, trazendo todos os requisitos pertinentes, cumprindo informar apenas que o relatório e os resultados foram apresentados ao Banco Mundial, destinatários da avaliação, conforme sugerido pelos autores na capa do documento. Recebeu conceito **A**, na meta-avaliação realizada.

No quinto e último Ponto de Verificação desse nível, (15) Meta-avaliação, apesar de serem encontrados indícios de atendimento a alguns dos padrões do Joint Committee (Utilidade, Viabilidade, Adequação e Precisão), não foram atingidos plenamente, considerando-se que a meta-avaliação não foi realizada. Julgado improcedente, recebeu o conceito **E**.

Conclusões E Recomendações

Os processos de avaliação, e, por conseguinte de meta-avaliação, devem ser reflexivos e críticos, sem perder de vista aqueles [processos] que podem ou devem ser considerados de

qualidade e que desempenham papel importante na construção do conhecimento sobre o objeto avaliado. No caso desse estudo meta-avaliativo, ficou demonstrado que a avaliação da educação superior não pode ser debatida isoladamente nas instituições de ensino superior, fazendo-se necessário o envolvimento de toda a sociedade, tornando-se ferramenta de autonomia das instituições e de todos os sujeitos que delas fazem parte.

Dos 14 Pontos de Verificação investigados, um, pertencente ao nível Subavaliações, foi descartado por não se adequar ao estudo meta-avaliativo proposto; nove foram considerados plenamente atendidos, sendo três em cada nível da lista; dois atingiram a maioria dos aspectos relevantes e ele relacionados em dois níveis diferentes: Fundamentos e Subavaliações; no nível Conclusões, um trata do ponto de forma desorganizada e pouco clara; outro tem elementos tratados de forma implícita, no nível Fundamentos; e, no nível Conclusões, um item perde ponto totalmente por não ter sido realizado.

Percebe-se, pelo exposto, que em cada nível verificado existem Pontos de Verificação que não foram considerados ou que não foram tratados de forma cuidadosa na avaliação, recebendo conceitos **C**, **D** e **E**. Porém, o atingimento total e/ou da maioria dos aspectos dos Pontos de Verificação que receberam, respectivamente, conceito **A** e **B**, totaliza 11, dos 14 investigados. Nesse caso, portanto, pode-se inferir que o relatório Avaliação da Educação Superior no Brasil: do Provão ao ENADE evidenciou uma avaliação considerada de boa qualidade, tendo em vista o atingimento da maioria dos Pontos de Verificação da Lista Chave de Verificação de Scriven de 2007, conforme demonstrado, contemplando o valor e o mérito preconizados pela meta-avaliação.

Entretanto, o estudo apontou que a utilização integral de todos os padrões do *Joint Committee* pelos autores, objetivando suscitar maior sustentação e clareza nas dimensões que se deseja atingir, possibilitaria acrescentar maior qualidade aos estudos avaliativos a serem realizados.

Referências

ADLER, C. S. *A avaliação de desempenho como processo essencialmente humano: a meta-avaliação da empresa X*. 2002. 131 f. Dissertação (Mestrado em Administração Pública)-Centro de Formação e Pesquisa, Fundação Getúlio Vargas, Rio de Janeiro, 2002. Disponível em: <<http://bibliotecadigital.fgv.br/dspace/handle/10438/3619>>. Acesso em: 11 abr. 2017.

ARROYO, D. M. P.; ROCHA, M. S. P. de M. L. da. Meta-avaliação de uma extensão universitária: estudo de caso. *Avaliação*, Campinas; Sorocaba, v. 15, n. 2, p. 135-161, jul. 2010. Disponível em: <<http://www.scielo.br/pdf/aval/v15n2/a08v15n2>>. Acesso em: 11 abr. 2017.

BALBIM, R. et al. *Meta-avaliação: estudos e proposições metodológicas a partir da avaliação de políticas de urbanização de assentamentos precários*. Brasília, DF: Instituto de Pesquisa Econômica Aplicada – IPEA, 2012. (Texto para discussão; n. 1704). Disponível em: <<https://www.econstor.eu/bitstream/10419/91249/1/718574915.pdf>>. Acesso em: 11 abr. 2017.

DAVIDSON, E. J. *Evaluation methodology basics*. Thousands Oaks, CA: Sage, 2005.

DAVOK, D. F. *Modelo de meta-avaliação de processos de avaliação da qualidade de cursos de graduação*. 2006. 274 f. Tese (Doutorado em Engenharia de Produção)-Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2006. Disponível em: <<https://repositorio.ufsc.br/bitstream/handle/123456789/88899/231045.pdf?sequence=1&isAllowed=y>>. Acesso em: 11 abr. 2017.

ELLIOT, L. G. Meta-avaliação: das abordagens às possibilidades de aplicação. Ensaio: aval. pol. públ. educ., Rio de Janeiro, v. 19, n. 73, out./dez. 2011.

ELLIOT, L. G. (Org.). *Instrumentos de avaliação e pesquisa: caminhos para construção e validação*. Rio de Janeiro: Wak, 2012.

FURTADO, J. P.; LAPERRIÈRE, H. *Parâmetros e paradigmas em meta-avaliação: uma revisão exploratória e reflexiva*. Ciência & Saúde Coletiva, Rio de Janeiro, v. 17, n. 3, mar. 2012. Disponível em: <http://www.scielo.org/scielo.php?pid=S1413-81232012000300016&script=sci_abstract&tlng=es>. Acesso em: 11 abr. 2017.

HEDLER, H. C.; TORRES, C. V. *Meta-avaliação de auditorias de natureza operacional do Tribunal de Contas da União*. Curitiba: RAC, 2009. Disponível em: <<http://www.scielo.br/pdf/rac/v13n3/v13n3a08>>. Acesso em: 11 abr. 2017.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. *The program evaluation standards: a guide for evaluators and evaluations users*. Los Angeles, CA: Sage, 2011. Parte 1: Padrões de Utilidade. Tradução Neuza Capelo; revisão técnica: Ligia Gomes Elliot.

PINTO, R. S. *Meta-avaliação: uma década do processo de avaliação institucional do SINAES*. 2015. 269 f. Tese (Doutorado em Administração)-Programa de Pós-Graduação em Administração da Universidade Federal de Santa Catarina, Florianópolis, 2015. Disponível em: <<https://repositorio.ufsc.br/xmlui/handle/123456789/136489>>. Acesso em: 11 abr. 2017.

PENNA FIRME, T.; LETICHEVSKY, A. C. *O desenvolvimento da capacidade de avaliação no século XXI: enfrentando o desafio através da meta-avaliação*. Revista Meta: Avaliação, Rio de Janeiro, v. 2, n. 5, 2010. Disponível em: <<http://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/79>>. Acesso em: 11 abr. 2017.

SCRIVEN, M. *Evaluation thesaurus*. 4. ed. Newbury Park, Ca: Sage, 1991.

SCRIVEN, M. *Evaluation checklists*. Michigan: Western Michigan University, 2007. Disponível em: <www.wmich.edu/evalctr/checklists>. Acesso em: 11 abr. 2017.

STUFFLEBEAM, D. L. The metaevaluation imperative. *American Journal of Evaluation*, Thousand Oaks, v. 22, n. 2, p. 183-209, 2001.

VERHINE, R. E.; DANTAS, L. M. V. *Avaliação da educação superior no Brasil: do Provão ao ENADE*. [S.l.], 2005. Disponível em: <<http://www.isp.ufba.br/avalia%C3%A7%C3%A3o%20da%20Ed%20Superior%20do%20Provao%20ao%20ENADE.pdf>>. Acesso em: 11 abr. 2017.

VERHINE, R. E.; DANTAS, L. M. V.; SOARES, J. F. *Do Provão ao ENADE: uma análise comparativa dos exames nacionais utilizados no Ensino Superior Brasileiro*. Ensaio: aval. pol. públ. educ., Rio de Janeiro, v. 14, n. 52, jul./set. 2006. Disponível em: <<http://www.scielo.br/pdf/ensaio/v14n52/a02v1452.pdf>> Acesso em: 11 abr. 2017.

WORTHEN, B. R.; SANDERS, J. R.; FITZPATRICK, J. L. *Avaliação de programas: concepções e práticas*. São Paulo: Gente, 2004.

Recebido em: 19/04/2018

Aceito para publicação em: 27/09/2018

Evaluation of Higher Education in Brazil – from the *Provão* to the National Student Proficiency Exam: a meta-evaluative study

Abstract

The study aimed to meta-evaluate the report presented to the World Bank in 2005, which addresses public policies for the evaluation of higher education implemented in Brazil. Scriven's Key Checklist (2007) was used as the meta-evaluative procedure, supported by Davidson's criteria (2005). As a result, the report has shown that there has been improvement in higher education assessments, but some points still need to be made. However, the full use of all Joint Committee standards by the authors would allow for higher quality evaluative studies to be carried out in the near future.

Keywords: Evaluation. Meta-evaluation. Evaluation standards. Key Evaluation Checklist.

Evaluación de la Educación Superior en Brasil – de la Prueba Grande (*Provão* en su original) al ENADE: un estudio metaevaluativo evaluativo

Resumen

El objetivo de este estudio fue metaevaluar el informe presentado al Banco Mundial en 2005, que se ocupa de las políticas públicas para la evaluación de la educación superior implementadas en Brasil. En el procedimiento metaevaluativo se utilizó la Lista Clave de Verificación de Scriven (2007), basada en los criterios establecidos por Davidson (2005). Como resultado, el informe mostró una mejora en las evaluaciones de la educación superior, pero algunos puntos todavía carecen de ciertos avances. Sin embargo, el estudio mostró que la plena utilización de todas las normas o patrones del Joint Committee, por parte de los autores, permitirá añadir mayor calidad a los estudios evaluativos que se lleven a cabo en el futuro.

Palabras clave: Evaluación. Metaevaluación. Normas o patrones de evaluación. Lista Clave de Verificación.