

MEC

Ministério da Educação

INEP

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

Avaliação de Cursos na Educação Superior: a função e a mecânica do Conceito Preliminar de Curso

Reynaldo Fernandes*

Elaine Toldo Pazello**

Thiago Miguel Sabino de Pereira Leitão***

Gabriela Miranda Moriconi****

* Presidente do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e professor da Universidade de São Paulo (USP). reynaldo.fernandes@inep.gov.br

** Diretora de Estudos Educacionais do Inep/MEC e professora da USP. elaine.pazello@inep.gov.br

*** Coordenador de Instrumentos e Medidas Educacionais do Inep/MEC e mestre em Economia pela Universidade do Porto, em Portugal. thiago.leitao@inep.gov.br

**** Coordenadora-Geral de Instrumentos e Medidas Educacionais do Inep/MEC e mestre em Administração Pública e Governo pela Escola de Administração de Empresas de São Paulo da Fundação Getulio Vargas (Eaesp/FGV). gabriela.moriconi@inep.gov.br

Brasília-DF
2009

© Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)
É permitida a reprodução total ou parcial desta publicação, desde que citada a fonte.

ASSESSORIA TÉCNICA DE EDITORAÇÃO E PUBLICAÇÕES

EDITOR EXECUTIVO

Jair Santana Moraes

NORMALIZAÇÃO BIBLIOGRÁFICA

Rosa dos Anjós Oliveira

REVISÃO

Antonio Bezerra Filho

DIAGRAMAÇÃO E ARTE-FINAL

Márcia Terezinha dos Reis

TIRAGEM

1.000 exemplares

EDITORIA

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)

Assessoria Técnica de Editoração e Publicações

Edifício-sede do Inep

SRTVS – Quadra 701, Lote 12, Bloco M – Ed. Dario Macedo, Térreo – Asa Sul

70340-909 – Brasília-DF

Fones: (61) 3799-3075 (61) 3799-3076 – Fax: (61) 3799-3079

editoria@inep.gov.br

DISTRIBUIÇÃO

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)

Edifício-sede do Inep

SRTVS – Quadra 701, Lote 12, Bloco M – Ed. Dario Macedo, Térreo – Asa Sul

70340-909 – Brasília-DF

Fone: (61) 3799-3060

publicacoes@inep.gov.br

<http://www.publicacoes.inep.gov.br>

ESTA PUBLICAÇÃO NÃO PODE SER VENDIDA. DISTRIBUIÇÃO GRATUITA.
A exatidão das informações e os conceitos e opiniões emitidos são de exclusiva
responsabilidade dos autores.

Dados Internacionais de Catalogação na Publicação (CIP)
Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

Avaliação de cursos na educação superior: a função e a mecânica do Conceito Preliminar de
Curso / Reynaldo Fernandes... [et al.]. – Brasília : Instituto Nacional de Estudos e Pesquisas
Educacionais Anísio Teixeira, 2009.

18 p. : il. – (Série Documental. Textos para Discussão, ISSN 1414-0640 ; 32)

1. Educação superior. 2. Avaliação de cursos. 3. Conceito preliminar de curso. I.
Fernandes, Reynaldo. II. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio
Teixeira.

CDU 378.146.004.12

Sumário

Avaliação de Cursos na Educação Superior: a função e a mecânica do Conceito Preliminar de Curso

RESUMO	5
INTRODUÇÃO	5
1 O SINAES E SEUS INDICADORES DE CURSOS	6
2 OBTENDO UM INDICADOR COMPOSTO DE QUALIDADE DE CURSOS	8
3 APLICAÇÃO DA METODOLOGIA	13
CONSIDERAÇÕES FINAIS	16
REFERÊNCIAS BIBLIOGRÁFICAS	18



Avaliação de Cursos na Educação Superior: a função e a mecânica do Conceito Preliminar de Curso

Reynaldo Fernandes

Elaine Toldo Pazello

Thiago Miguel Sabino de Pereira Leitão

Gabriela Miranda Moriconi

RESUMO

Este artigo analisa a finalidade, a qualidade e a coerência dos diversos indicadores de qualidade de cursos produzidos no âmbito do Sinaes. A ênfase é dada ao Conceito Preliminar de Cursos (CPC), o qual se pretende um indicador da contribuição do curso para a formação dos alunos e tem por finalidade orientar e racionalizar a avaliação in loco. Por ser um indicador composto, o CPC tem recebido a crítica de juntar coisas muito diferentes. Este texto esclarece que todos os índices que compõem o CPC podem ser vistos como medidas imperfeitas do que se pretende medir: a contribuição do curso para a formação dos seus alunos. Uma metodologia para combinar as diversas medidas imperfeitas existentes foi desenvolvida e aplicada para os cursos de graduação, de modo a obter uma medida com menor erro. Os resultados mostram que o CPC se aproxima muito do indicador “ideal” encontrado. No entanto, o estudo mostra que é possível encontrar um indicador tão simples quanto o CPC e mais correlacionado com o indicador “ideal”, com a vantagem de dispensar a inclusão da nota dos ingressantes no conceito Enade e, ainda assim, combater o incentivo ao gaming.

INTRODUÇÃO

A avaliação dos cursos de graduação no País se dá no âmbito do Sistema Nacional

de Avaliação da Educação Superior (Sinaes) e objetiva aferir tanto a qualidade acadêmica do profissional egresso quanto a contribuição do curso para sua formação.

A qualidade média dos profissionais egressos do curso, por exemplo, é uma informação que auxilia empregadores e potenciais estudantes em suas escolhas. Como a boa qualidade do profissional dá prestígio ao curso e orienta a demanda, a disponibilidade dessa informação incentiva as instituições a melhorarem seu desempenho.

Porém, a qualidade média do egresso não é um indicador direto do trabalho da instituição. Um curso que recebe alunos com formação deficiente pode realizar um excelente trabalho e, ainda assim, não ter seus ex-alunos entre os melhores profissionais. Para os gestores, e também para a sociedade e órgãos reguladores, é importante saber quanto o curso contribui para a formação dos alunos.

A comparação dos cursos com base em indicadores de qualidade é, em si, valiosa para quem necessita tomar decisões relacionadas ao assunto, sejam elas de ordem pessoal ou gerencial. No entanto, a mera comparação pode não ser suficiente. Existe na sociedade uma demanda por critérios de qualidade, em especial a identificação de cursos que não atendem a um nível mínimo de qualidade; aliás, é função do regulador definir a qualidade mínima para o funcionamento dos cursos.

Assim, o Sinaes propõe estipular critérios de qualidade com base na contribuição do curso para a formação dos alunos.

O objetivo do presente texto é analisar a finalidade, a qualidade e a coerência dos diversos indicadores de qualidade de cursos produzidos no âmbito do Sinaes. A ênfase é dada ao Conceito Preliminar de Cursos (CPC), o qual se pretende um indicador da contribuição do curso para a formação dos alunos e tem por finalidade orientar e racionalizar a avaliação *in loco*.

A legislação vigente prevê que todos os cursos sejam visitados a cada três anos, o que, dado o grande número de cursos, exigiria uma estrutura muito maior que a disponível. Além do mais, a experiência brasileira mostra que os resultados de avaliações *in loco* não guardam muita relação com indicadores objetivos da qualidade dos cursos. Ao dispensar da visita os cursos com indicador 3, 4 e 5, o CPC viabiliza as avaliações *in loco* e, por servir de guia para os avaliadores, induz uma maior coerência entre essas avaliações e os indicadores objetivos de qualidade.

Por ser um indicador composto, o CPC tem recebido a crítica de juntar coisas muito diferentes e, assim, ser um indicador sem sentido. Neste texto mostramos que todos os índices que compõem o CPC podem ser vistos como medidas imperfeitas de uma mesma coisa: a contribuição do curso para a formação dos seus alunos. Quando se possuem diversas medidas imperfeitas de uma mesma coisa é possível juntá-las e obter uma medida com menor erro. O problema com o CPC é que ele não utilizou, na combinação inicial de seus componentes, nenhum procedimento explícito de minimização da variância do erro, o que é agora proposto.

Este texto é organizado da seguinte forma: a seção 1 analisa os indicadores de cursos do Sinaes, destacando o papel do CPC para a coerência do sistema; a seção 2

apresenta uma forma de ponderar distintos indicadores para se obter um indicador mais eficiente da qualidade do curso; na seção 3, finalmente, esta metodologia é aplicada para os cursos de educação superior, e o resultado é comparado com o atual CPC.

1 O SINAES E SEUS INDICADORES DE CURSOS

Os principais instrumentos de avaliação de cursos do Sinaes são o Exame Nacional de Desempenho dos Estudantes (Enade) e a avaliação *in loco*, conduzida por especialistas que visitam os cursos superiores.

O Enade aplica provas para alunos concluintes e ingressantes. Os dois grupos de estudantes fazem a mesma prova, que é composta por 30 questões de formação específica e 10 de formação geral. Além da prova, o Enade conta com um questionário por meio do qual o estudante fornece informações sobre sua condição socioeconômica e sobre sua percepção acerca da qualidade de diferentes aspectos do curso. Na avaliação *in loco*, os avaliadores dão seu parecer em relação às condições de ensino em três distintas dimensões: organização didático-pedagógica, perfil do corpo docente e técnico-administrativo e instalações físicas. Além das informações produzidas por esses dois instrumentos, o Sinaes conta, para produção de indicadores, com as informações disponibilizadas pelo censo e pelo cadastro de docentes da educação superior.

A pontuação média dos concluintes de determinado curso na prova do Enade fornece um indicador da qualidade acadêmica média dos profissionais egressos desse curso. Considerando apenas as questões de caráter específico, esse indicador é similar ao obtido pelo antigo Exame Nacional de Cursos (ENC), mais conhecido como Provão.¹ Nesse sentido,

¹ Para uma comparação entre o Enade e o Provão, ver Verhine, Dantas e Soares (2006).

as restrições feitas ao Enade são as mesmas que se faziam ao Provão: a) o que é passível de ser medido em uma prova não representa os conhecimentos e as habilidades requeridas para o exercício profissional; e b) a prova é uma medida imperfeita dos objetivos restritos que ela se propõe avaliar.

Em que medida a intensidade e a gravidade desses problemas podem comprometer a aferição da qualidade acadêmica dos egressos de um curso não é algo fácil de avaliar. Tais problemas devem, provavelmente, variar de área para área e, para uma determinada área, de uma edição para outra do Enade.² Seria de se esperar que, em cursos com grande conteúdo prático – como, por exemplo, odontologia –, a primeira restrição fosse mais importante do que em cursos mais teóricos – como, por exemplo, matemática. Já a segunda restrição diz respeito a erros de medida, que surgem em virtude de a nota da prova não depender apenas dos conhecimentos e habilidades que ela procura testar. O resultado na prova depende também da motivação e preparação específica dos estudantes para realizar o exame, das condições da aplicação, da sorte, etc. Em estudo para o ensino fundamental na Carolina do Norte, Kane e Staiger (2001, 2002) notaram que os resultados de exames padronizados são medidas sujeitas a muito ruído, particularmente entre as pequenas escolas.

O fato é que não possuímos muitos estudos que nos auxiliem no esclarecimento dessas questões – que avaliem, por exemplo, o poder preditivo do Enade ou do Provão sobre o desempenho profissional futuro dos ex-alunos. Portanto, muito do que se discute a respeito desse assunto acaba sendo baseado em percepções individuais. Vale ressaltar que, quando se necessita de um critério mais objetivo para julgar a competência de

profissionais para, por exemplo, contratar pessoal no setor público ou conceder certificação profissional, a aplicação de uma prova é ainda o expediente mais utilizado.

O Enade possibilita, também, calcular uma medida direta da contribuição da escola para a formação acadêmica de seus estudantes. Essa medida é dada pelo Indicador de Diferença entre os Desempenhos Observado e Esperado (IDD), que calcula a diferença entre a nota obtida pelos concluintes e a nota que seria esperada (baseada na nota dos ingressantes). Ainda que informativo, o IDD possui seus inconvenientes.

O IDD é, na verdade, uma medida da contribuição do curso para o desempenho de seus estudantes no Enade – ou seja, todas as restrições que se possam fazer ao Enade estão presentes no IDD. Aliás, no que se refere aos erros de medida da prova, o IDD agrava o problema, pois, além das notas dos concluintes, ele se utiliza das notas dos ingressantes.³ O IDD, também, não permite estabelecer critérios de qualidade, o que seria de fundamental importância para o regulador, que necessita decidir quais os cursos que atendem às condições de funcionamento. Como o Enade não possui uma escala predeterminada, fica impossível estabelecer critérios de qualidade em medidas dele derivadas; essas medidas servem apenas para comparar cursos entre si. Por fim, o IDD é, provavelmente, uma medida enviesada contra os cursos que, por sua qualidade, atraem bons alunos. O IDD é o resíduo de uma regressão, entre cursos, da nota média dos concluintes sobre a nota média dos ingressantes e sobre algumas variáveis de controle. Se a qualidade do curso está positivamente correlacionada com as notas dos ingressantes, o IDD subestima a qualidade dos cursos que recebem alunos com bom desempenho.⁴

² Como o Enade não utiliza escalas pré-definidas e nem pré-testagem de itens, a qualidade da prova, em termos de cobertura e poder de discriminação, pode variar entre diferentes edições.

³ Se o IDD fosse a simples diferença entre as médias de notas de concluintes e ingressantes e a variância do erro de medida fosse igual para os dois grupos, a variância do erro de medida no IDD seria o dobro do que na nota média dos concluintes.

⁴ Em razão do caráter amostral da prova e devido ao fato de a menor parte dos estudantes terminarem o curso em quatro anos, é difícil a utilização de notas individuais, em vez de agregadas por curso, para a estimação do IDD. Para uma avaliação dos modelos de valor agregado, ver, por exemplo, McCaffrey et al. (2004).

Outra medida da qualidade do ensino oferecido aos estudantes é dada pelo Conceito de Curso, que é obtido a partir de visitas *in loco*. Nesse caso, os conceitos atribuídos são baseados em critério em que o indicador 3 significa que o curso atende às condições mínimas de funcionamento. A vantagem da avaliação *in loco* é poder observar características específicas dos cursos que não podem ser detectadas com indicadores padronizados. As dificuldades se referem à dimensão da avaliação e à comparabilidade e coerência dos conceitos. Como mencionado anteriormente, avaliar todos os mais de 22 mil cursos a cada três anos exigiria uma estrutura muitas vezes maior do que a hoje existente. Por outro lado, o conceito dos cursos pode variar de acordo com diferentes avaliadores. Introduzir um roteiro rígido que reduza a influência da subjetividade dos avaliadores pode não funcionar. A experiência brasileira mostra que os conceitos tendem a ser altos, mesmo quando indicadores objetivos se mostram insatisfatórios.

É nessa perspectiva que surge o Conceito Preliminar de Cursos (CPC). Obtido pela combinação de indicadores de qualidade já existentes, o CPC tem a função de orientar e racionalizar as avaliações *in loco*. Se maior que 2, o curso pode dispensar a visita e, assim, o CPC torna-se automaticamente o Conceito de Curso. Para os cursos visitados (por solicitação ou por ter o CPC inferior a 3), o Conceito de Curso é dado pelos avaliadores, mas, caso divirja do CPC, é necessário uma justificação sólida dos motivos. Com isso espera-se eliminar a aparente contradição entre indicadores objetivos e os da avaliação *in loco*. Assim, o CPC é um indicador intermediário que viabiliza e dá consistência ao Conceito de Curso.

Uma possibilidade teria sido utilizar o IDD para cumprir esse papel. Entretanto, os estudos realizados pelo Inep mostraram que o IDD tem maior variabilidade do que os demais indicadores do Enade. Além do mais, se boas escolas atraem bons alunos, o IDD tende a ser enviesado contra as escolas que atraem bons alunos.

Então, se utilizássemos a nota dos concluintes como medida da qualidade do curso estaríamos beneficiando cursos que atraem bons alunos por outros motivos que não a qualidade, como, por exemplo, por serem gratuitos. Por outro lado, se utilizássemos o IDD estaríamos prejudicando as escolas que, por serem boas, atraem bons alunos.

Nesse caso, o ideal seria combinar os indicadores de modo a minimizar a variância do erro. Foi com base nessas considerações que a ponderação entre IDD e nota do Enade foi proposta. Já a inclusão das variáveis de insumo foi realizada com base no poder explicativo da variável sobre o IDD. A idéia é de que tais variáveis tendem a ser mais estáveis do que as notas do Enade e, assim, sua inclusão visa reduzir a variância da medida de qualidade. A principal deficiência do CPC, no entanto, reside no fato de ele não contar com nenhum mecanismo explícito que justifique as ponderações utilizadas. Na seção seguinte é apresentado um procedimento para se obter tal ponderação.

2 OBTENDO UM INDICADOR COMPOSTO DE QUALIDADE DE CURSOS

Seja C_{kj} a pontuação no Enade obtida pelo aluno k quando concluinte no curso j e I_{kj} a nota desse aluno quando ingressou no curso. Então, a medida de valor agregado para esse aluno durante o período em que frequentou o curso j é dada por $C_{kj} - I_{kj}$. Vamos admitir também que o modelo que determina o valor agregado é dado por (1).

$$c_{kj} - i_{kj} = q_j + bi_{kj} + x'_{kj}\phi + v_{kj}, \quad (1)$$

$$v_{kj} = w_{kj} + z_j.$$

Em (1) a letra minúscula das variáveis significa que elas estão expressas como desvio da média, de forma que nenhuma constante se faz necessária. A variável q_j é um índice da qualidade da escola j , x_{kj} é um vetor de variáveis de controle expressando,

por exemplo, o ambiente familiar dos estudantes e v_{kj} é o termo referente ao distúrbio aleatório, onde $E[v_{kj}|q_j] = 0$. O fato de i_{kj} entrar no lado direito da equação significa que agregar valor depende do desempenho dos ingressantes. Se $b < 0$ ($b > 0$), é mais difícil (mais fácil) agregar valor para estudantes de alto desempenho. O termo w_{kj} é um distúrbio aleatório que afeta o aluno independentemente do curso em que ele se encontra; já o termo z_j é um distúrbio aleatório que afeta todos os estudantes do curso j (pode refletir um erro de procedimento do aplicador do exame ou algum evento que afetou a motivação e a concentração dos estudantes para realizar o teste). Note-se que os dois tipos de distúrbio devem levar a erros na medida da proficiência dos alunos, porém, enquanto o primeiro pode ser reduzido com o aumento de indivíduos que fazem a prova, o segundo independe do número de observações.

Admite-se, ainda, que $E[w_{kj}] = E[z_j] = 0$, $Var[w_{kj}] = \sigma_w^2$, $Var[z_j] = \sigma_z^2$ e $Cov[w_{kj}, z_j] = 0$. O fato de q_j entrar aditivamente na equação (1) significa que não existem escolas especializadas para determinados perfis de alunos. Assim, se os alunos ingressantes na escola j fossem colocados na escola h (no lugar dos ingressantes da escola h), eles teriam um ganho esperado de valor adicionado dado por $q_h - q_j$, independentemente de i_{kj} . Agregando (1) por cursos, temos:

$$c_j - i_j = q_j + bi_j + x_j'\phi + v_j, \quad (2)$$

$$v_j = w_j + z_j,$$

$$Var[v_j] = \frac{\sigma_w^2}{n_j} + \sigma_z^2.$$

Em (2) observamos, como já colocado, que o primeiro termo da variância do erro se reduz com o aumento do número de alunos testados (n_j), mas não o segundo. Assim, a

variância de v_j é heterocedástica, mas, por simplificação, vamos desconsiderar esse fato e admitir que $Var[v_j] = \sigma_v^2$. Note-se que, se os ingressantes de um curso no período t possuem desempenho equivalente ao dos concluintes quando estes ingressaram no curso, podemos considerar c_j e i_j em (2) como a pontuação média de concluintes e ingressantes da mesma edição do Enade.

Sejam \hat{b} e $\hat{\phi}$ estimadores de b e ϕ , então o Indicador de Diferença entre os Desempenhos Observado e Esperado seria dado por (3).

$$\begin{aligned} idd_j &= c_j - (1 + \hat{b})i_j - x_j'\hat{\phi} = \\ &= q_j + (b - \hat{b})i_j + x_j'(\phi - \hat{\phi}) + v_j. \end{aligned} \quad (3)$$

Para completar a especificação do modelo, admita-se que o desempenho médio dos ingressantes seja uma função da qualidade do curso e do tipo de instituição. Definindo P_j como uma variável *dummy* que assume valor 1 quando o curso for oferecido por uma instituição pública, temos:

$$i_j = \gamma q_j + \eta p_j + u_j. \quad (4)$$

A expressão (4) pode ser justificada de duas maneiras: a primeira é que cursos melhores recebem mais candidatos e, assim, podem realizar uma melhor seleção; a segunda é que o exame de ingressantes do Enade é realizado ao final do primeiro ano e, desse modo, algum valor agregado já foi realizado pelo curso. A *dummy* pública procura capturar o fato de as escolas públicas receberem, em média, melhores alunos tanto em função de sua qualidade como em função da gratuidade. O termo u_j reflete todos os demais fatores que afetam a qualidade do

ingressante e não dizem respeito à qualidade do curso, como, por exemplo, a facilidade de acesso. Admite-se que $E[u_j|q_j] = E[u_j] = 0$ e $Var[u_j] = \sigma_u^2$.⁵

Com base em (2), (3) e (4), podemos observar que a nota média dos ingressantes, a nota média dos concluintes e o IDD são de fato medidas imperfeitas da qualidade do curso, que diferem por um termo de erro e por um fator de escala.

$$c_j = Z_1 \gamma q_j + (1+b)u_j + (1+b)\eta p_j + x_j' \phi + v_j, \quad (2')$$

$$idd_j = Z_2 \gamma q_j + (b-\hat{b})u_j + (b-\hat{b})\eta p_j + x_j'(\phi - \hat{\phi}) + v_j, \quad (3')$$

$$i_j = \gamma q_j + \eta p_j + u_j, \quad (4)$$

$$Z_1 = \frac{1}{\gamma} + (1+b),$$

$$Z_2 = \frac{1}{\gamma} + (b-\hat{b}).$$

Caso Z_1 , Z_2 e η fossem conhecidos, seria possível eliminar as diferenças de escala entre os diferentes indicadores de qualidade do curso, o que é dado por (5).

$$c_j^* = \frac{c_j - (1+\hat{b})\eta p_j - x_j' \hat{\phi}}{Z_1} = \gamma q_j + e_{1j}, \quad (5)$$

$$idd_j^* = \frac{idd_j}{Z_2} = \gamma q_j + e_{2j},$$

$$i_j^* = i_j - \eta p_j = \gamma q_j + e_{3j},$$

$$e_{1j} = \frac{(1+b)u_j + (b-\hat{b})\eta p_j + x_j'(\phi - \hat{\phi}) + v_j}{Z_1},$$

$$e_{2j} = \frac{(b-\hat{b})u_j + (b-\hat{b})\eta p_j + x_j'(\phi - \hat{\phi}) + v_j}{Z_2},$$

$$e_{3j} = u_j.$$

E, assim, podemos definir um indicador w dado por:

$$w_j = \alpha c_j^* + \beta idd_j^* + (1-\alpha-\beta)i_j^* = \gamma q_j + e_j, \quad (6)$$

$$e_j = \alpha e_{1j} + \beta e_{2j} + (1-\alpha-\beta)e_{3j}.$$

O objetivo passa a ser, então, encontrar α^* e β^* que definem w_j^* , tal que $Var[w_j^*|q_j]$ seja a menor possível. Pode-se mostrar que:

$$Var[w_j|q_j] = \alpha^2 Var[e_{1j}] + \beta^2 Var[e_{2j}] + (1-\alpha-\beta)^2 Var[e_{3j}] + 2\alpha\beta Cov[e_{1j}, e_{2j}] + 2\alpha(1-\alpha-\beta) Cov[e_{1j}, e_{3j}] + 2\beta(1-\alpha-\beta) Cov[e_{2j}, e_{3j}] \quad (7)$$

E as condições de primeira ordem para minimizar (7) são dadas por (8) e (9).

⁵ Novamente, como erros de medida da prova estão capturados pelo termo erro, a variância de u_j deve ser heterocedástica, o que não está sendo levado em conta.

$$\alpha^* = \alpha' - \beta^* \frac{X_3}{X_{13}}, \quad (8)$$

$$\beta^* = \beta' - \alpha^* \frac{X_3}{X_{23}}, \quad (9)$$

$$\alpha' = \frac{Var[e_{3j}] - Cov[e_{1j}, e_{3j}]}{X_{13}},$$

$$\beta' = \frac{Var[e_{3j}] - Cov[e_{2j}, e_{3j}]}{X_{23}},$$

$$X_3 = Var[e_{3j}] + Cov[e_{1j}, e_{2j}] - Cov[e_{1j}, e_{3j}] - Cov[e_{2j}, e_{3j}],$$

$$X_{13} = Var[e_{1j}] + Var[e_{3j}] - 2Cov[e_{1j}, e_{3j}],$$

$$X_{23} = Var[e_{2j}] + Var[e_{3j}] - 2Cov[e_{2j}, e_{3j}].$$

O termo α' em (8) é o α que minimiza a variância, quando apenas c_j^* e i_j^* são considerados. Já β' é o coeficiente ótimo quando apenas idd_j^* e i_j^* são levados em conta. Admitindo que $Cov[q_j, e_{hj}] = 0$ ($h = 1, 2$ e 3), podemos obter estimadores para α^* e β^* .⁶

$$\alpha^* = \frac{4Var[i_j^*] - Var[c_j^* + i_j^*] + Var[c_j^* - i_j^*]}{4X_{13}}, \quad (10)$$

$$\beta^* = \frac{4Var[i_j^*] - Var[idd_j^* + i_j^*] + Var[idd_j^* - i_j^*]}{4X_{23}}, \quad (11)$$

$$X_{13} = Var[c_j^* - i_j^*], \quad (12)$$

$$X_{23} = Var[idd_j^* - i_j^*], \quad (13)$$

$$X_3 = \frac{Var[c_j^* - i_j^*] + Var[idd_j^* - i_j^*] - Var[c_j^* - idd_j^*]}{2}. \quad (14)$$

Note-se que podemos obter uma estimativa para cada uma das variâncias contidas em (10)-(14) e, assim, obtermos $\hat{\alpha}^*$ e $\hat{\beta}^*$. No entanto, para implementar o procedimento acima, necessitamos de estimativas para Z_1 , Z_2 e η .

Seja L um vetor de variáveis de insumos e processos educacionais que afetem a qualidade do curso, de modo que:

$$\gamma q_j = l'_j \pi + \theta_j. \quad (15)$$

Em (15), θ_j é um termo aleatório que inclui todos os fatores que afetam a qualidade e não estão incluídos em l_j . Admite-se que

$$E[\theta_j | l_j] = E[u_j | l_j] = E[v_j | l_j] = 0. \text{ De (4) e (15) temos:}$$

$$i_j = l'_j \pi + \theta_j + \eta p_j + u_j. \quad (16)$$

Então, podemos obter $\hat{\pi}$ e $\hat{\eta}$ por MQO e, utilizando $l'_j \hat{\pi}$ como instrumento, obter \hat{Z}_1 e \hat{Z}_2 . Definindo $\bar{c}_j = c_j - (1 + \hat{b}) \hat{\eta} p_j - x'_j \hat{\phi}$ e $\lambda_j = \eta p_j + u_j = \hat{\eta} p_j + \hat{u}_j$, podemos reescrever (2') e (3') como:

⁶ Como $Cov[v_j, q_j] = Cov[u_j, q_j] = 0$, as hipóteses adicionais $Cov[x_j, q_j] = Cov[p_j, q_j] = 0$ seriam suficientes para garantir que $Cov[q_j, e_{hj}] = 0$. Essas hipóteses, no entanto, podem não ser verificadas e, assim, algum viés estará presente nas estimativas de α^* e β^* . O tamanho desse viés depende do viés de \hat{b} e $\hat{\phi}$. Se \hat{b} e $\hat{\phi}$ forem estimativas não enviesadas de b e ϕ , então $\hat{\alpha}^*$ e $\hat{\beta}^*$ também serão não enviesados.

$$\bar{c}_j = Z_1 \gamma q_j + (1+b) \hat{u}_j + (b-\hat{b}) \hat{\eta} p_j + x'_j (\phi - \hat{\phi}) + v_j, \quad (2'')$$

$$idd_j = Z_2 \gamma q_j + (b-\hat{b}) \hat{u}_j + (b-\hat{b}) \hat{\eta} p_j + x'_j (\phi - \hat{\phi}) + v_j. \quad (3'')$$

Multiplicando (2'') e (3'') por $l'_j \hat{\pi}_j$ e extraindo o somatório entre todos os cursos, temos:

$$\sum_{j=1}^N \bar{c}_j l'_j \hat{\pi} = Z_1 \sum_{j=1}^N \gamma q_j l'_j \hat{\pi} + G_1, \quad (17)$$

$$\sum_{j=1}^N idd_j l'_j \hat{\pi} = Z_2 \sum_{j=1}^N \gamma q_j l'_j \hat{\pi} + G_2, \quad (18)$$

$$G_1 = (1+b) \sum_{j=1}^N \hat{u}_j l'_j \hat{\pi} + (b-\hat{b}) \sum_{j=1}^N \hat{\eta} p_j l'_j \hat{\pi} + \left(\sum_{j=1}^N x_j l'_j \hat{\pi} \right) (\phi - \hat{\phi}) + \sum_{j=1}^N v_j l'_j \hat{\pi},$$

$$G_2 = (b-\hat{b}) \sum_{j=1}^N \hat{u}_j l'_j \hat{\pi} + (b-\hat{b}) \sum_{j=1}^N \hat{\eta} p_j l'_j \hat{\pi} + \left(\sum_{j=1}^N x_j l'_j \hat{\pi} \right) (\phi - \hat{\phi}) + \sum_{j=1}^N v_j l'_j \hat{\pi}.$$

Admitindo que $G_1 \cong 0$ e $G_2 \cong 0$, o estimador para $Z = \frac{Z_1}{Z_2}$ é dado por:

$$\hat{Z} = \frac{\sum_{j=1}^N \bar{c}_j l'_j \hat{\pi}_j}{\sum_{j=1}^N idd_j l'_j \hat{\pi}_j}. \quad (19)$$

De posse de \hat{Z} e sabendo que $Z_1 - Z_2 = (1+\hat{b})$, obtemos \hat{Z}_1 e \hat{Z}_2 . Note-se que o procedimento acima para obter \hat{Z} é idêntico a regressir, por MQO, \bar{c}_j e idd_j contra $l'_j \hat{\pi}$, obter as estimativas de Z_1 e Z_2 e, então, extrair a razão. Pode-se observar que tanto o primeiro como o último termo de G_1 e G_2 possuem esperança igual a zero. O problema de viés pode surgir do fato de as correlações entre l'_j e p_j e entre l'_j e x_j serem diferentes de zero. Nesse caso, o tamanho do viés vai depender do viés de \hat{b} e $\hat{\phi}$. Novamente, se \hat{b} e $\hat{\phi}$ forem estimativas não enviesadas de b e ϕ , então \hat{Z} também será não enviesado.⁷

Note-se que nosso indicador w_j^* é uma combinação de c_j^* , idd_j^* e i_j^* , portanto nenhuma variável de insumo ou processo educacional o compõe. As variáveis de insumos e processos entraram apenas como instrumento para a estimação de Z_1 , Z_2 . No entanto, como essas variáveis possuem informações sobre a qualidade do curso, elas, em princípio, poderiam integrar o indicador, obtendo-se assim uma medida de qualidade mais eficiente. Definindo w_j^{**} como

$$w_j^{**} = \tau l'_j \pi + (1-\tau) w_j^*, \quad (20)$$

$$w_j^* = \gamma q_j + e_j^*,$$

$$w_j^* = l'_j \pi + \theta_j + e_j^*.$$

Substituindo (15) em (20), temos:

$$w_j^{**} = \gamma q_j + \tau \theta_j + (1-\tau) e_j^*. \quad (21)$$

⁷ Ver nota 6.

Admitindo que q_j , θ_j e e_j^* são independentes, o τ^* que minimiza a variância do erro em (21) é dado por:

$$\tau^* = \frac{Var[e_j^*]}{Var[\theta_j] + Var[e_j^*]} \quad (22)$$

O problema em (22) é que não dispomos de um estimador para a $Var[e_j^*]$.

Ou seja, embora exista um τ^* que minimiza a variância do erro em (21) nós não conseguimos identificá-lo. Note-se que

$\frac{Var[\theta_j] + Var[e_j^*]}{Var[w_j^*]}$ é igual a $(1 - R^2)$ da

regressão, por MQO, de w_j^* em l_j' . Então, podemos escrever:

$$(1 - R^2)\tau^* = \frac{Var[e_j^*]}{Var[w_j^*]} \quad (23)$$

Uma regra de bolso seria considerar que $\tau^* = R^2$. Assim, quanto melhor o ajuste da regressão de w_j^* em l_j' , maior o peso dado às variáveis de insumos e processos. Note-se que essa regra implica que e_j^* explica, no máximo, 25% da variância de w_j^* (isso quando o R^2 é igual a 0,5). Então, o indicador adotado é dado por (24), onde $\tilde{\pi}$ e R^2 são obtidos pela regressão de w_j^* em l_j' .

$$w_j^{**} = R^2 l_j' \tilde{\pi} + (1 - R^2)w_j^* \quad (24)$$

O procedimento descrito nesta seção possui o inconveniente de ser de difícil comunicação com o público interessado. Em primeiro lugar, o indicador w_j^{**} não é composto

de índices de conhecimento público, como nota do Enade e IDD, mas, antes, de transformações deles, c_j^* , idd_j^* e i_j^* . Em segundo lugar, os pesos tendem a variar de área para área, o que pode ser difícil para explicar. Então, seria conveniente encontrar um indicador que: a) os pesos não variassem entre as áreas; b) fosse composto pelos índices originais (C_j , IDD_j e I_j), além dos insumos; e c) possuisse alta correlação com w_j^{**} .

Para obter tal indicador, todas as variáveis (w_j^{**} , C_j , IDD_j , I_j e indicadores de insumos e processos) foram padronizadas para variar entre 0 e 5, conforme a fórmula do conceito Enade. Essa padronização é realizada para cada uma das áreas. Após tal padronização, estimou-se uma regressão por MQO de w_j^{**} em C_j , IDD_j , I_j e indicadores de insumos e processos. Os coeficientes dessa regressão, devidamente reponderados para somar 1, foram utilizados como pesos. O resultado desse procedimento é mostrado na seção seguinte.

3 APLICAÇÃO DA METODOLOGIA

A metodologia desenvolvida para obter um indicador de qualidade de curso foi aplicada à totalidade dos cursos de ensino superior avaliados. Os dados utilizados são provenientes dos Enades 2005, 2006 e 2007, de onde foram extraídas as variáveis referentes ao desempenho, às condições socioeconômicas e à percepção dos estudantes sobre a qualidade de insumos e processos. Também foram utilizados dados provenientes do Cadastro de Docentes de 2007.

No cálculo do CPC, as variáveis de insumos e processos utilizadas foram: o percentual de docentes com título mínimo de doutor; o percentual de docentes com regime de trabalho integral ou parcial; a percepção dos alunos em relação a um

aspecto da infraestrutura do curso; e a percepção dos alunos em relação a um aspecto da organização pedagógica do curso.⁸ Essas variáveis foram incluídas por serem, entre todas testadas, as que se mostraram correlacionadas com o IDD. Neste texto incluímos, além das variáveis acima, o percentual de docentes com título mínimo de mestre.

O primeiro passo foi estimar o IDD para cada curso de graduação de cada IES. A estimação foi realizada de forma separada para cada área avaliada com base no Enade, sendo considerados todos os cursos participantes que cumpram os seguintes requisitos:

- o número de ingressantes participantes deve ser de no mínimo 10 e deve representar pelo menos 20% dos ingressantes totais do curso;
- o número de concluintes participantes deve ser de no mínimo 10 e deve representar pelo menos 20% dos concluintes totais do curso.

O cálculo do IDD consiste em estimar por MQO a equação $c_j = (1+b)i_j + x_j'\hat{\phi} + e_j$, onde c_j e i_j excluem as provas zeradas. As variáveis de controle incluídas são a razão entre ingressantes e concluintes do curso e a proporção de alunos do curso cujo pai tem nível superior de escolaridade. Na estimação, o resíduo (e_j) é composto da qualidade (q_j) mais um erro. No entanto, como já destacado, se a nota dos ingressantes é correlacionada com a

qualidade do curso, $(1+\hat{b})$ será enviesado. Para se ter uma estimativa melhor de $(1+b)$, foram inseridos nessa regressão insumos que explicam parte da qualidade. Portanto, a regressão estimada foi $c_j = (1+b)i_j + x_j'\hat{\phi} + l_j'\varphi + e_j$, onde l_j' é composto pelas variáveis de insumos e processos descritas anteriormente.

Nessa estimação, o coeficiente dos ingressantes não se mostrou significativo a 10% para as seguintes áreas: Fonoaudiologia, Teatro, Engenharia V, Arquivologia e Tecnologia em Agroindústria. Deste modo, optou-se por não imputar valor ao IDD dessas áreas e, assim, excluí-las do estudo.

Para se obter a estimativa para Z_1 e Z_2 , o primeiro passo é obtermos $l_j'\hat{\pi}$. Essa estimativa é gerada de forma separada para cada área avaliada, sendo que foram considerados para a estimação de $l_j'\pi$ apenas os insumos que demonstraram ter um efeito positivo e significativo a 10% sobre a nota dos ingressantes. Além das áreas já excluídas na etapa anterior, foram excluídas nessa etapa as seguintes áreas por não apresentarem nenhuma das variáveis de insumos significativas: Medicina, Terapia Ocupacional, Engenharia VIII e Tecnologia em Radiologia.

Para obter \hat{Z} , o procedimento foi regredir, por MQO, \bar{c}_j e idd_j contra $l_j'\hat{\pi}$, obter as estimativas de Z_1 e Z_2 e, então, extrair a razão. Foram mantidas no estudo apenas as áreas cujos coeficientes Z_1 e Z_2 foram positivos e significativos a 10%. Desse modo, foram excluídas as seguintes áreas: Física, Farmácia, Filosofia, Música, Zootecnia, Ciências Sociais, Engenharia III, Engenharia IV, Engenharia VI, Engenharia VII,

⁸ A pergunta sobre infraestrutura utilizada é a seguinte: "Aulas práticas: os equipamentos disponíveis são suficientes para todos os alunos?". Respostas possíveis: 'A' = Sim, em todas elas. / 'B' = Sim, na maior parte delas. / 'C' = Sim, mas apenas na metade delas. / 'D' = Sim, mas em menos da metade delas. / 'E' = Não, em nenhuma. A partir desta pergunta construiu-se uma variável binária que é igual a '1' caso o aluno tenha respondido os itens 'A' ou 'B' e igual a zero em caso contrário. A pergunta sobre a área pedagógica é: "Os planos de ensino contêm todos os seguintes aspectos: objetivos; procedimentos de ensino e avaliação; conteúdos e bibliografia da disciplina". 'A' = Sim, todos contêm. / 'B' = Sim, a maior parte contêm. / 'C' = Sim, mas apenas cerca da metade contêm. / 'D' = Sim, mas apenas menos da metade contêm. / 'E' = Não, nenhum contêm. A partir desta pergunta construiu-se uma variável binária que é igual a '1' caso o aluno tenha respondido 'A' e igual a zero em caso contrário.

Biblioteconomia, Secretariado e Normal Superior.

Para se obter w_j^{**} , as provas zeradas foram desconsideradas. Entretanto, para construção do indicador-síntese, os zeros foram considerados tanto em c_j como em i_j . O primeiro passo na construção desse indicador foi transformar todas as variáveis utilizadas (w_j^{**} , C , I , IDD e variáveis de insumos e processos) em valores padronizados entre 0 e 5. Essa padronização foi feita primeiramente transformando as variáveis em desvios padronizados, conforme $X^{padr} = \frac{X - \bar{X}}{\sigma_X}$,

onde X representa a variável a ser padronizada, \bar{X} representa a sua média e σ_X significa o seu respectivo desvio padrão.

Após essa transformação, os valores foram padronizados em escala de 0 a 5 de

$$\text{acordo com } X^{padr(0a5)} = 5 \frac{X^{padr} - X^{padr_inf}}{X^{padr_sup} - X^{padr_inf}},$$

onde X^{padr_inf} representa o menor valor padronizado para a respectiva variável e X^{padr_sup} representa o maior valor padronizado para a respectiva variável. Para excluir a influência dos valores muito discrepantes (*outliers*), os valores de X^{padr_inf} e X^{padr_sup} utilizados foram restringidos aos limites de -3 a 3 respectivamente.

O último passo foi regredir, por MQO, o indicador w_j^{**} contra os insumos, C , I e IDD , todos padronizados entre 0 e 5. Dois modelos foram considerados: M1, que inclui a constante, e M2, que não inclui constante. Os pesos obtidos encontram-se na Tabela 1. Já a Tabela 2 apresenta a correlação de ordem entre os diversos indicadores.

Tabela 1 – Comparação dos pesos dos componentes dos indicadores

	Doutor	Mestre	Regime	Infra	Pedag	C	I	IDD	Enade
M1	0,146	0,044	0,010	0,048	0,034	0,104	0,206	0,408	0,000
M2	0,117	0,062	0,022	0,069	0,044	0,061	0,199	0,426	0,000
CPC	0,117	0,000	0,071	0,031	0,081	0,000	0,000	0,300	0,400

Tabela 2 – Teste de Correlação de Rank de Spearman

	W^*	w^{**}	M1	M2	CPC ⁽¹⁾	C	I	IDD
w^*	1,000							
w^{**}	0,979	1,000						
M1	0,850	0,872	1,000					
M2	0,841	0,866	0,997	1,000				
CPC ⁽¹⁾	0,800	0,828	0,929	0,936	1,000			
C	0,806	0,799	0,863	0,845	0,821	1,000		
I	0,648	0,638	0,647	0,632	0,472	0,557	1,000	
IDD	0,627	0,618	0,742	0,748	0,815	0,687	0,126	1,000

⁽¹⁾ Os CPCs foram recalculados para cada curso utilizando o novo IDD desenvolvido nesse estudo, e não representam os CPCs já calculados e divulgados pelo Inep.

Podemos observar pela Tabela 1 que o peso das variáveis de insumos e processos não se altera muito entre os dois modelos (M1 e M2) e o atual CPC. Ele fica em torno de 30%. Entretanto, ocorre uma mudança de pesos entre as variáveis de insumos e processos. Em relação ao CPC, nos modelos aqui estimados as variáveis de regime e organização pedagógica perdem importância, aumentando o peso da infraestrutura. Por fim, a variável proporção de docentes com título mínimo de mestre passa a ser significativa, com peso de 4,4% e 6,2% de M1 e M2, respectivamente.

A Tabela 2 mostra que nosso indicador “ideal”, w_j^{**} , é extremamente correlacionado com w_j^* , de modo que a inclusão das variáveis de insumos e processos não trouxe grande contribuição ao indicador. As correlações entre o indicador “ideal” e M1 e M2 foram, também, bastante elevadas, indicando que o preço a pagar, em termos de precisão, para se obter uma maior simplicidade, não é tão alto. Na realidade, as correlações entre w_j^{**} e qualquer dos demais indicadores considerados se mostraram elevadas. Assim, alguém poderia questionar se vale a pena substituir o atual CPC por M1, dado que a correlação entre eles é de 0,93.

Em primeiro lugar, é preciso reconhecer que, embora a correlação seja elevada, para alguns cursos a mudança de posição, por se adotar um ou outro indicador, pode ser razoável. Mas existe um motivo mais importante para se adotar M1 em vez do atual CPC. Para o cálculo do conceito Enade, a nota dos ingressantes é considerada,⁹ e isto tem sido motivo de críticas. Como mencionado anteriormente, se o interesse é obter uma medida da qualidade acadêmica do profissional egresso, então a inclusão dos ingressantes no conceito Enade é, no mínimo, estranha.

O motivo dessa inclusão foi no sentido de buscar evitar a prática de *gaming* por parte das escolas. Note-se que, se apenas os conceitos Enade e IDD fossem divulgados e se o conceito Enade incluísse apenas os concluintes, as escolas poderiam incentivar os ingressantes a terem um mau desempenho. Isso não afetaria o conceito Enade e elevaria, artificialmente, o IDD. Como em M1 a nota do ingressante entra explícita e positivamente, o conceito Enade poderia se restringir aos concluintes. Assim, teríamos duas medidas principais: a nota dos concluintes no Enade e M1. A primeira seria um indicador da qualidade acadêmica do profissional egresso e a segunda, um indicador da contribuição do curso para sua formação. As demais medidas (nota do ingressante, IDD e insumos) seriam apenas indicadores intermediários para a medida de valor adicionado.

CONSIDERAÇÕES FINAIS

O presente texto teve o objetivo de analisar a finalidade, a qualidade e a coerência dos diversos indicadores de qualidade de cursos produzidos no âmbito do Sinaes. A ênfase foi dada ao CPC, o qual se pretende um indicador da contribuição do curso para formação dos alunos. Vimos que o CPC é fundamental para viabilizar a avaliação de cursos no Sinaes, pois seria impossível visitar todos os cursos a cada três anos. Ao dispensar da visita os cursos com maiores pontuações, o CPC viabiliza as avaliações *in loco*.

Além disso, o CPC induz a uma maior coerência entre as avaliações *in loco* e os indicadores objetivos de qualidade. Se maior que 2, o curso pode dispensar a visita e, assim, o CPC torna-se o Conceito de Curso. Para os cursos visitados (por solicitação ou por ter o CPC inferior a 3), o Conceito de Curso é dado

⁹ A nota do componente específico de concluintes entra com peso de 0,60 e de ingressantes com peso de 0,15, enquanto a nota de formação geral de concluintes e ingressantes entra com peso de 0,25.

pelos avaliadores. Mas, caso diverja do CPC, é necessária uma justificação dos motivos.

Por ser um indicador composto, o CPC tem recebido a crítica de juntar coisas muito diferentes e, assim, ser um indicador sem sentido. Neste texto mostramos que todos os índices que compõem o CPC podem ser vistos como medidas imperfeitas de uma mesma coisa: a contribuição do curso para a formação dos seus alunos. Uma metodologia para se combinar diversos indicadores de resultados, insumos e processos foi desenvolvida e aplicada para os cursos de educação superior. Os resultados mostram que o CPC se aproxima muito do que denominamos indicador “ideal”. No entanto,

o estudo mostra que é possível encontrar um indicador tão simples quanto o CPC e mais correlacionado com o indicador “ideal”. Tal indicador possui a vantagem de dispensar a inclusão da nota dos ingressantes no conceito Enade e, ainda assim, combater o incentivo ao *gaming*.

A conclusão é de que deveriam existir apenas dois indicadores principais de qualidade extraídos do Enade: a nota dos concluintes e uma medida de valor adicionado (um novo CPC). Os demais indicadores seriam apenas intermediários para a medida de valor adicionado.

REFERÊNCIAS BIBLIOGRÁFICAS

KANE, T. J.; STAIGER D. O. Improving school accountability measures. *NBER Working Paper*, n. 8156, Mar. 2001.

_____. The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*, v. 16, n. 4, p. 91-114, Fall 2002.

McCAFFREY, D. F. et al. Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, v. 29, n. 1, Value-Added Assessment Special Issue, p. 67-101, 2004.

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: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 14, n. 52, p. 291-310, jul./set. 2006.





