

Avaliação Qualitativa e Quantitativa: Uma Metodologia de Integração

João Carlos Mello / Fabiana Leta
Artur Fernandes / Marcos Vaz
Maria Helena / Myriam Barbejat

RESUMO

Avaliar é fundamental em qualquer processo. Especialmente no processo educativo, onde a avaliação pode ser um instrumento para alavancar mudanças no ensino superior, objetivando a melhoria da qualidade, e reordenar recursos. O processo de aprendizagem é individual, apesar de na grande maioria ser tratado como coletivo. Frequentemente são realizadas "avaliações" pautadas em opiniões vagas ou imprecisas, ou em dados estatísticos sem nenhuma análise prévia. Neste trabalho, por acreditarmos que os esforços envolvidos com a Educação devem ser aplicados na direção certa e produtiva, encaramos novos procedimentos que busquem um ponto de equilíbrio na unificação da avaliação qualitativa com a quantitativa.

Palavras-chave: Avaliação – Metodologias de avaliação – Decisões multicritério – Análise de dados.

João Carlos Mello

Eng^o Mecânico

Fabiana Leta

Eng^o Mecânico

Artur Fernandes

Eng^o Eletricista

Marcos Vaz

Eng^o Eletricista

Maria Helena Campos

Eng^o Mecânico

Myriam Barbejat

Eng^o Mecânico

Professores da Escola de Engenharia da Universidade Federal Fluminense - RJ

Membros do Grupo de Estudos sobre Ensino Tecnológico – GESET

1. Introdução

A questão da avaliação do ensino vem sendo bastante explorada por profissionais de diversas correntes de pensamento e metodologia. Em geral, os procedimentos em uso neste campo são calcados em conceitos essencialmente qualitativos, fortemente dependentes da opinião do avaliador. Quando usadas versões quantitativas, normalmente observa-se a aplicação de índices brutos, sem análise de causa-conseqüência e sem estudos comparativos. Segundo Boclin (1999): "a avaliação deve ser quantitativa e comparativa", sugerindo que a visão apresentada ao longo deste artigo é bastante pertinente e atual,

por incorporar características qualitativas e de opinião dentro de um método quantitativo/comparativo.

Nas Escolas de Engenharia, e demais unidades de ensino tecnológico, a palavra avaliação sofre de um certo paradoxo. Considerando que, em quase todas as habilitações, futuros profissionais são ensinados a fazer avaliações, entre outras, da confiabilidade de uma peça; do valor de uma empresa; da eficiência de um processo, verifica-se que, numa ótica genérica, a avaliação de sua própria formação ou não é feita, ou quando realizada incorre em equívocos. Estes enganos são os mesmos de quem, ao contrário dos profissionais da área tecnológica, não tem formação fortemente quantitativa. Também no contexto de ensino, são comuns alterações curriculares, experiências didáticas, mudanças na forma de ingresso, baseadas apenas em considerações fortuitas, do tipo: "eu acho", "no lugar tal fizeram assim", "agora está reprovando menos", e outras semelhantes.

Deste modo, torna-se essencial, principalmente no ambiente onde se aprende a raciocinar com método e objetividade, introduzir técnicas de avaliação que possam ser usadas sem influência de aspectos subjetivos. No entanto, essas técnicas não devem incorrer no erro inverso, de serem meras contas que não incorporem o conhecimento de especialistas no assunto e do método.

Quando se tenta fazer alguma avaliação quantitativa, alguns erros clássicos são cometidos:

- Avaliação de situações complexas usando um único critério (por exemplo, índice de reprovação).
- Tentativa de avaliação absoluta, não levando em conta que avaliar é essencialmente um processo de comparação.
- Falhas ao transformar avaliações qualitativas em escalas comparativas, com utilização de associações rígidas sem margem de ponderações.
- Avaliação pontual no tempo, não levando em conta situações passadas, ou o que pode vir futuramente a acontecer.

Na âmbito da Engenharia, o eterno, e globalizado debate sobre Cálculo I exemplifica bem esses enganos. É comum a afirmação: "O professor X foi ruim, porque reprovou 80%". Evidentemente essa afirmação, isoladamente, não possui um grande significado, mas a sua habitual contestação é igualmente questionável: "O índice de reprovação é apenas 50%, porque vários alunos abandonaram e não podem ser levados em conta". Neste caso, cada afirmação leva em conta um fator diferente, e por isso mostra realidades diferentes. Uma melhor aproximação da realidade deve levar em conta: (1) a aprovação em relação ao total de alunos, ou em relação aos que não abandonaram; (2) qual a condição dos alunos quando começaram o curso; (3) que aproveitamento tiveram depois do curso; (4) opiniões qualitativas transformadas corretamente numa escala quantitativa; (5) comparações com outras turmas e até com outras instituições semelhantes.

Uma outra situação, igualmente merecedora de destaque, decorre da seguinte afirmação: "A turma B de Física II foi pior que a A, porque reprovou 60% contra 40%". Esta afirmativa já faz uma comparação, mas continua usando apenas um critério, e não leva em conta outras variáveis, como as notas na Física e no Cálculo anteriores, na Física seguinte, no prosseguimento do curso, entre outras.

Vê-se portanto que a avaliação é um processo bem mais complexo que a simples comparação de índices.

2. O Auxílio Multicritério à Decisão

Este recente ramo do conhecimento é uma poderosa ferramenta quando se quer apreciar diversas variáveis (critérios) na avaliação, e alguns dos seus métodos permitem a conversão de escalas qualitativas em quantitativas. Moldando-se a diferentes correntes de pensamento, os métodos multicritério permitem montar estruturas de preferência sem uma indicação definitiva, ou agregar todos os critérios analisados em um único critério que sintetize todas as informações dos vários critérios. À primeira característica costuma-se chamar de Escola Francesa, e de Escola Americana à segunda. Mas, em qualquer dos casos, importa ressaltar que os métodos não fornecem um resultado automático e singular. Eles sempre dependem fortemente da opinião de especialistas no assunto (muitas vezes nem tão especialistas,

mas possuidores do poder de decisão) para estabelecer preferências entre alternativas, relações hierárquicas, pesos de critérios, etc. Embora esta situação seja vantajosa ao permitir incorporar opiniões, ela tem a desvantagem de ser fortemente arbitrária, conduzindo a resultados sob condição de serem falseados por opiniões extremadas do responsável pela decisão.

Na sua forma mais popularizada, os métodos multicritério fazem, para cada alternativa em análise, uma soma ponderada dos valores obtidos nos vários pesos. Ou seja, para cada alternativa obtém-se um valor $F = \sum a_i c_i$ sendo c_i o valor obtido no critério i e a_i o peso desse critério. As alternativas são então classificadas por F .

É exatamente na determinação dos pesos que se incorpora a opinião do especialista/decisor e, portanto, onde reside a parcela de arbitrariedade do método. Entretanto, existem formas para reduzir essa arbitrariedade, onde adquire notável importância a análise de sensibilidade. Para a determinação dos pesos são permitidos vários métodos para ajudar o decisor a fazer uma escolha mais objetiva, porém sempre refletindo as suas opiniões. Entre eles podem ser citados o UTA (UTilité Additive), AHP (Análise Hierárquica de Processos) e, principalmente, MACBETH (Bana e Costa, 1997), cujo uso será posteriormente apresentado e defendido.

Resta, ainda, a dificuldade de transformar os critérios qualitativos em escalas

quantitativas cardinais. Nos métodos oriundos da psicologia essa transformação é realizada através de questionários que pedem para classificar cada alternativa numa escala de 0 a 100. Esse método exige que o avaliador tenha uma idéia extremamente precisa da escala de valores, o que a experiência mostra nem sempre acontecer. Além disso, o método citado exige uma classificação absoluta, o que é irreal. Qualquer atribuição de valor é comparativa. Perguntar se o professor foi ótimo, bom ou medíocre; se o conteúdo apresentado contribui muito, pouco ou nada para a formação, só tem sentido quando existe uma referência para poder fazer a comparação. Os métodos oriundos da engenharia dão mais ênfase à comparação. Em especial a comparação entre diferenças de alternativas mostra-se um método com resultados bastante positivos. Pode-se perguntar se o professor de, por exemplo, Cálculo I apresentou maior ou menor disponibilidade que o professor de Física I e se essa diferença foi grande ou pequena. Com base nessa informação, o já citado MACBETH sugere, usando programação linear, uma escala cardinal de valores, permitindo a opção do decisor por fazer pequenos ajustes, sem violar as suas próprias escolhas anteriores.

De qualquer forma, todos esses métodos apenas auxiliam o decisor a quantificar as suas opções, que sempre influenciam decisivamente o critério síntese encontrado ao final.

3. A Análise Envoltória de Dados (DEA)

Surgido formalmente no final da década de 70, este método (Charnes, Cooper, Rhodes, 1978) foi desenvolvido para avaliar a eficiência de unidades produtivas que consomem múltiplos recursos e têm vários produtos gerados. Na versão inicial do método, essas unidades produtivas eram supostas serem capazes de tomar decisão sobre o nível utilizado de recursos (entradas) e a quantidade fabricada de cada produto (saídas) e por isso receberam a sigla de DMU – Decision Making Units.

A grande vantagem da DEA é que permite que cada unidade busque a eficiência de forma distinta, de tal maneira que atenda às suas especificidades, dependendo, assim, de sua vocação. Em termos matemáticos, a DEA calcula a eficiência fazendo a razão entre uma soma ponderada das saídas (*output*) e uma soma ponderada das entradas (*input*). Mas, ao contrário do modelo multicritério, em DEA os pesos usados nas ponderações não surgem da opinião determinativa de algum decisor. São, ao contrário, calculados segundo um critério objetivo. Consiste esse critério em determinar os pesos da maneira mais favorável a cada DMU. Isso é feito maximizando-se a eficiência de cada DMU, com as restrições de que os pesos de uma DMU aplicados às outras não podem gerar eficiências superiores a 100%. Em termos formais, tem-se o seguinte problema de programação matemática:

$$EF_k = \text{Máx} \frac{\sum_j u_{jk} \cdot y_{jk}}{\sum_i v_{ik} \cdot x_{ik}}$$

com as restrições:

$$\frac{\sum_j u_{jk} \cdot y_{jk}}{\sum_i v_{ik} \cdot x_{ik}} \leq 1, \forall k$$

$$u_{jk}, v_{ik} \geq 0$$

Onde EF_k representa a eficiência da unidade k ; os y_{jk} significam as saídas (output) da DMU k ; os x_{ik} são as entradas (input) da mesma DMU; e os u_{jk} e v_{ik} definem os pesos usados na ponderação.

Embora este não seja um problema linear, é facilmente linearizável, o que permite o uso de vasta gama de programas existentes para programação linear.

4. Uso Integrado de DEA e Análise Multicritério

Como acabou de ser visto, os modelos DEA são mais objetivos que os modelos multicritério, reduzindo a influência subjetiva das opiniões de especialistas e, portanto, sendo menos passíveis de manipulação. Por outro lado, a estrutura matemática do modelo faz com que, freqüentemente, uma unidade seja considerada eficiente ao serem atribuídos pesos nulos a quase todos os coeficientes. O problema dessa situação é que vários

critérios acabam sendo desconsiderados na forma final, acarretando desta maneira uma avaliação incompleta e uma errônea ordenação das unidades. No entanto, existem técnicas adicionais para minorar esse problema e que, ainda, podem permitir um certo grau de incorporação das opiniões do decisor, fazendo uma mescla (mix) de objetividade e subjetividade (Araya et al, 1999).

Com esta nova forma de abordagem, pode-se inserir os pesos sugeridos pela análise multicritério no modelo DEA, reduzindo assim, a quantidade de unidades que se tornariam 100% eficientes, uma vez que elas deixariam de ser avaliadas exclusivamente pelos critérios em que são melhores ou mais benéficos.

O modelo MACBETH, adotado com este propósito, sugere uma faixa de variação dos pesos. Cabe ao decisor escolher qual peso, no interior da faixa, utilizar, ou mesmo como utilizar estes pesos.

5. Exemplo de Utilização

Para uma melhor compreensão da técnica apresentada, segue-se um exemplo onde mostramos sua utilização.

Considere-se a seguinte situação fictícia: aplicar a metodologia desenvolvida ao avaliar 12 cursos, da área tecnológica, cujos alunos ingressam na universidade por meio de vestibular, realizando provas de mesmo conteúdo. Os cursos escolhidos são:

- | | |
|------------------|-------------|
| 1. Engenharia A | – campus I |
| 2. Engenharia B | – campus II |
| 3. Engenharia C | – campus I |
| 4. Engenharia D1 | – campus I |
| 5. Engenharia D2 | – campus II |
| 6. Engenharia E | – campus I |
| 7. Engenharia F1 | – campus I |
| 8. Engenharia F2 | – campus II |
| 9. Engenharia G | – campus I |
| 10. Exatas H | – campus I |
| 11. Exatas I | – campus I |
| 12. Exatas J | – campus I |

O objetivo do problema é determinar quais os cursos são mais eficientes. Tome-se por critérios, segundo os quais se deseja medir esta eficiência, os relatados a seguir:

- 1- **Nota de ingresso** – sabendo-se que o mecanismo de ingresso é o mesmo para todas os cursos, será utilizada como dado de entrada (*input*), para cada curso, a nota média obtida no vestibular pelos alunos.
- 2- **Número de ingressantes** – será utilizado o número bruto de alunos ingressantes.
- 3- **Grau de satisfação** – este parâmetro será obtido a partir da opção de aprovação do aluno, ou seja, será considerado que se o aluno ingressou na universidade em sua primeira opção, haverá um grau de satisfação máximo, que no exemplo foi definido com o valor dez.

4- **Tempo de acesso à universidade** – será observado o tempo médio despendido pelos alunos em seus deslocamentos para chegarem à universidade, definindo-se com valor zero o menor tempo gasto e 100 o maior tempo. A obtenção destes valores se dá através da análise dos endereços residenciais dos alunos, informados na inscrição do vestibular, quando comparados com o endereço do campus onde o aluno estuda.

5- **Renda familiar** – será calculada a média de renda familiar dos alunos por curso. A renda familiar poderá ocasionar produtividades escolares diferenciadas dos alunos por curso. Considerar-se-á 100 a maior renda familiar média. Estes dados são obtidos em questionários respondidos pelos candidatos ao vestibular na época de sua inscrição.

Aos critérios acima citados são acrescidos os seguintes:

- 1- **Número de formados** – é o número de alunos formados por curso avaliado.
- 2- **Avaliação de saída** – dado referente ao resultado de avaliações realizadas por entidades externas ou ao índice de empregabilidade dos alunos, ou seja, o número de alunos empregados e/ou realizando pós-graduação dividido pelo número de formados. O índice de empregabilidade é determinado considerando-se um período de até quatro meses após a formatura.

No modelo DEA, os dois últimos critérios serão utilizados como *output* e os anteriores como *input*. Na Tabela 1 são apresentados os dados utilizados no exemplo. Ressalta-se que estes dados são fictícios. Embora, ao serem gerados, foram tomados como referência alguns cursos que realmente existem.

De forma sintética: (1) o tempo de acesso foi normalizado numa escala de 0 a 100, a partir de dados obtidos em horas; (2) a renda familiar também é normalizada a partir de valores em reais; (3) a satisfação é padronizada (0 a 10); (4) os números de ingressantes, de formados e a nota de ingresso são números absolutos; e (5) a avaliação externa de saída é percentual.

Tabela 1 – Dados Utilizados no Modelo para o Exemplo

Curso	Nota de ingresso	Número de ingressantes	Satisfação	Tempo de Acesso	Renda familiar	Formados	Avaliação externa
Eng. A	15	20	15	50	49	2	40
Eng. B	18	30	20	25	38	7	100
Eng. C	30	30	70	100	78	15	70
Eng. D1	60	30	95	73	100	22	90
Eng. D2	70	30	100	71	98	15	100
Eng. E	32	40	50	98	73	5	75
Eng. F1	31	30	68	97	75	13	70
Eng. F2	30	30	30	0	69	8	60
Eng. G	20	30	28	10	55	6	68
Exatas H	21	60	20	60	32	7	100
Exatas I	20	40	20	65	30	3	100
Exatas J	72	40	100	72	90	20	80

A Tabela 2 apresenta os resultados obtidos ao se inserirem os dados anteriores no modelo DEA, considerado o modelo isoladamente.

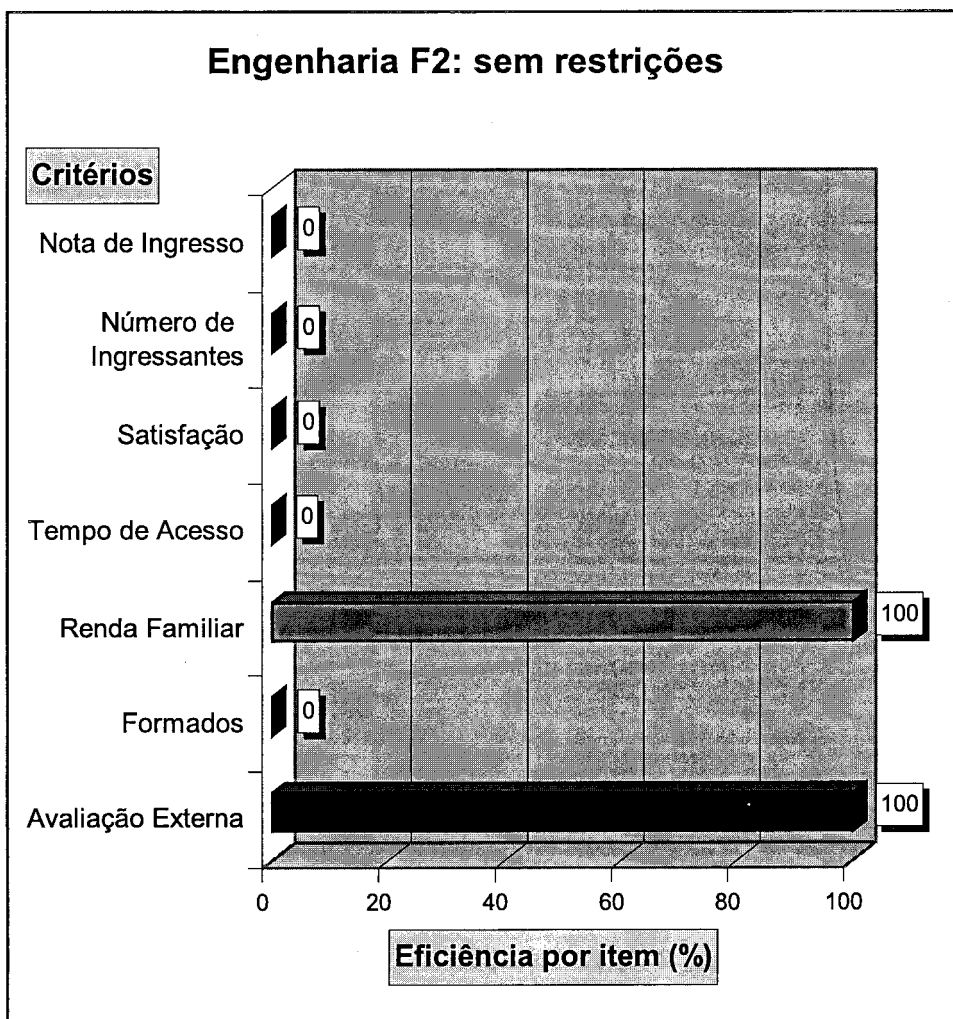
Tabela 2 – Resultados Obtidos pelo Modelo DEA

Curso	Eficiência %
Eng. F2	100
Eng. G	100
Exatas H	100
Exatas I	100
Eng. D1	100
Eng. C	100
Eng. B	100
Eng. D2	100
Exatas J	100
Eng. F1	90,67
Eng. A	60
Eng. E	56,25

Observe-se que quase todos os cursos foram considerados eficientes, devido ao fato de o modelo avaliar cada unidade segundo o critério que lhe seja mais favorável. Assim, por exemplo, o resultado iso-

lado para Engenharia F2, apresentado a seguir no gráfico 1, demonstra quais critérios (*inputs e outputs*) formam percentualmente uma situação mais vantajosa para o curso em análise.

Gráfico 1 – Resultado isolado para Engenharia F2



A seguir pode ser visto o resultado final com a aplicação integral do modelo proposto. Nele são verificados os efeitos da ação conjunta dos procedimentos definidos pelos métodos

multicritério e DEA. Pode ser claramente percebida a diferença muito mais acentuada quanto às eficiências dos cursos e, conseqüentemente, à seletividade do método.

Tabela 3 – Resultados Obtidos no Modelo Integrado

Curso	Eficiência %
Eng. B	100
Exatas H	90,69
Exatas I	76,96
Eng. G	65,32
Eng. F2	58,61
Eng. C	49,05
Eng. F1	48,10
Eng. D1	47,12
Eng. D2	43,01
Eng. A	40,79
Exatas J	38,98
Eng. E	37,95

Ainda, como efeito da utilização do modelo integrado proposto neste trabalho, podemos ver os gráficos 2 e 3, referentes ao ponto de vista isolado da En-

genharia B, onde é demonstrado o efeito das restrições impostas pelo modelo completo quando comparado ao modelo DEA isoladamente:

Gráfico 2 – Resultado isolado para Engenharia B sem restrições

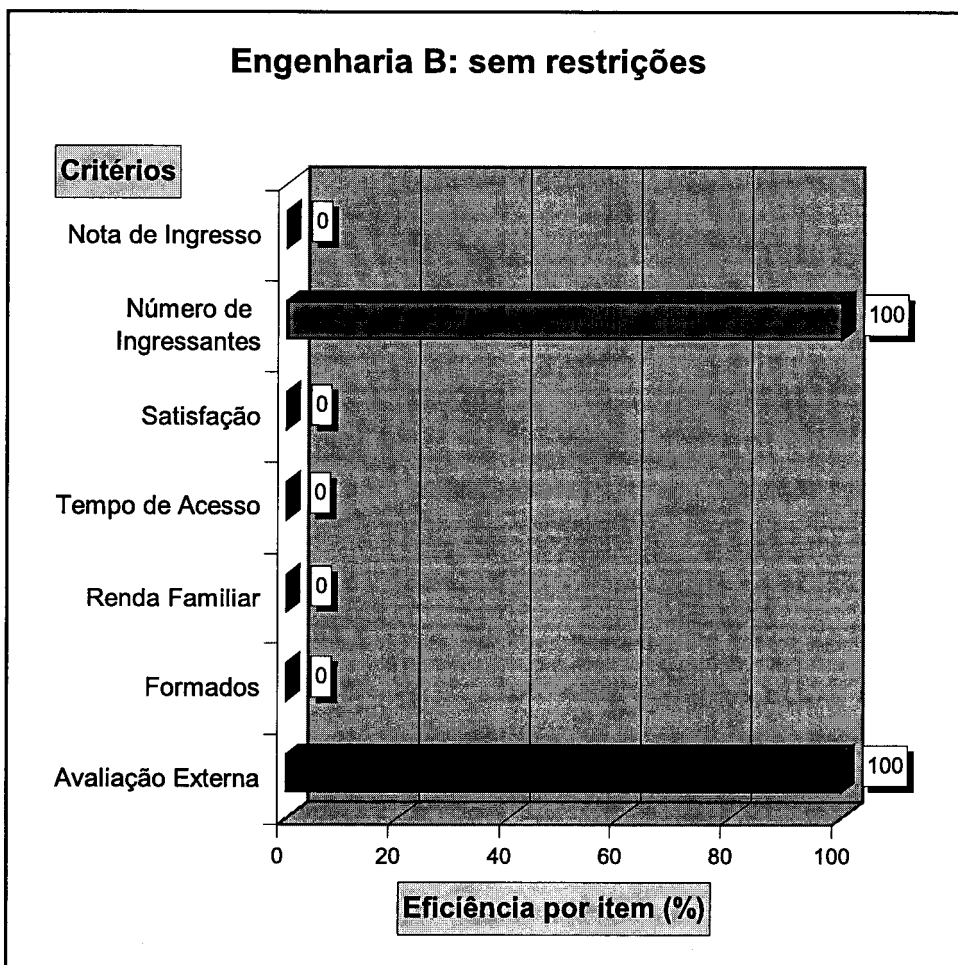
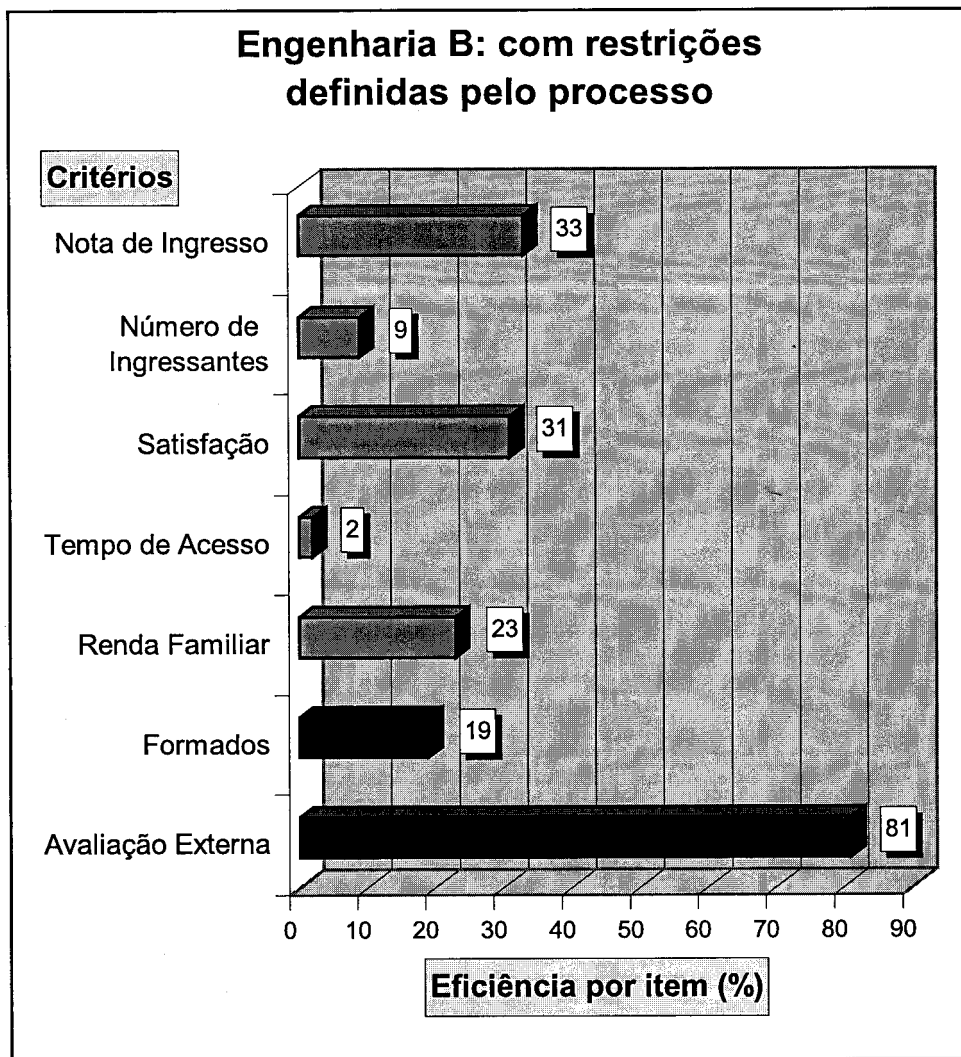


Gráfico 3 – Resultado isolado para Engenharia B
após aplicação do modelo integrado

6. Conclusões

O efeito observado pela aplicação do método proposto sobre dados normalizados e que, por ação do método DEA, não apresentam uma influência excessiva da vontade direcionada do decisor, demonstra a junção produtiva das visões quantitativa e qualitativa nos procedimentos de avaliação. Esta integração é garantida pela hierarquização das características sujeitas à avaliação (entradas e saídas), com certa

liberdade de intervenção pelo avaliador, e pela aplicação de processos existentes e conhecidos.

Saliente-se que a metodologia apresentada, com o uso da DEA, é interessante nos casos em que se apresentam muitas unidades para serem avaliadas, quando comparadas ao número de entradas e saídas disponíveis para análise.

ABSTRACT

To evaluate is essential in all processes, specially in the educational one, where the evaluation can be an instrument to reach changes in high education level, aiming the improvement of quality and rearrange resources. The learning process is individual, in spite of the big majority it is treated as a group process. Frequently "evaluations" based in uncertain and unsure opinions, or in statistics data without any previous analysis are done. In this article, as we believe that the efforts related to Education must be applied in the right and productive direction, we face new procedures that search for a balanced point in the unification of the quantity with the quality evaluation.

Keywords: Evaluation – Evaluation methodology – Multicriteria decisions – Data analysis.

RESUMEN

Evaluar es fundamental en cualquier proceso. Especialmente en el proceso educativo, donde la evaluación puede ser un instrumento de trampolín para cambios en la enseñanza superior, objetivando la mejora de la calidad y reordenando recursos. El proceso de aprendizaje es individual, a pesar de que en su gran mayoría se trate como colectivo. Frecuentemente se realizan "evaluaciones" pautadas en opiniones vagas o imprecisas, o en datos estadísticos sin ningún análisis previo. En este trabajo, por creer que los esfuerzos relacionados con la Educación se deben aplicar en una dirección cierta y productiva, encaramos nuevos procedimientos que busquen un punto de equilibrio en la unificación de la evaluación cualitativa con la cuantitativa.

Palabras-clave: Evaluación – Metodologías de evaluación – Decisiones multicriterio – Análisis de datos.

Referências Bibliográficas

- BANA E COSTA, C.A., VANSNICK, J.C. *Applications of the MACBETH approach in the framework of an additive aggregation model. Journal of Multi-Criteria Decision Analysis*, v. 6, p.107-14, 1997.
- BANA E COSTA, C.A., VANSNICK, J.C. *Cardinal Value Measurement with Macbeth*. Lisboa: CEG-IST, 1999. Research Paper nº 13.
- BARBA-ROMERO, S., POMEROL, J.C. *Decisiones multicriterio*, colección de Economía. [Madrid?] Universidad de Alcalá, 1997.
- BEASLEY, J-E. *Comparing University Departments*, OMEGA International. London, *Journal of Management Science*, v.18, n. 2, p.171-83, 1990.
- BOCLIN, R. Indicadores de desempenho: novas estratégias da Educação Superior. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 7, n. 24, p. 299-308, jul./set. 1999.
- CHARNES, A., COOPER, W.W., RHODES, E. *Measuring the efficiency of decision-making units. European Journal of Operational Research*, v. 2, n.6, p. 429-44, Nov. 1978.
- COELLI, T., RAO, P.D.S., BATTESE, G.E. *An Introduction to efficiency and productivity analysis*. Boston: Kluwer Academic Publishers, 1998.
- COOK, W.D., KRESS, M., SEIFORD, L. Data envelopment analysis in the presence of both quantitative and qualitative factors. *Journal of the Operational Research Society*, v.47, n.7 p. 945-53, July 1996.
- GONZALEZ-ARAYA, M.C., LINS, M.P.E., GOMES, L.F.A.M. A integração entre a análise de envoltória de dados e o apoio multicritério à decisão: uma revisão. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 31, 1999, Juiz de Fora, MG. *Anais...* [s.n.t.]
- LINS, M.P.E., MEZA, L.A. *Análise envoltória de dados e perspectivas de integração no ambiente do apoio à decisão*. Rio de Janeiro: COPPE/UFRJ, 2000.
- MELLO, J.C.C.B.S., LINS, M.P.E., MELLO, M.H.C.S. Avaliação da eficiência do ensino de cálculo utilizando a análise de envoltória de dados. In: CONGRESSO DA APDIO, 9, 2000, Setúbal (Port). [s.n.t.]
- MELLO, M.H.C.S., MELLO, J.C.C.B.S., LINS, M.P.E. Teaching calculus using computational techniques - quantitative analysis of experience in Universidade Federal Fluminense. INTERNATIONAL CONFERENCE ON ENGINEERING AND COMPUTER EDUCATION - ICECE, São Paulo, 2000.
- ROY, B., BOUYSSOU, D. *Aide multicritère à la décision: méthode et cas*. economica. Paris: Economica, c1993. 695 p. (Collection gestion. Serie: Production et techniques quantitatives appliques a la gestion)