

Sobre a Dificuldade da Realização de Estudos Avaliando o Desempenho de Estudantes que Desenvolvem Diferentes Currículos: Comparando o Incomparável?

Com respeito ao artigo "Avaliação do aproveitamento do aluno em Nefrologia: estudo comparativo entre o método expositivo tradicional e o método de aprendizado baseado em problemas", do professor Ivan de Mello Araújo, da Faculdade de Medicina de Marília – FAMEMA, e das professoras Rita de Cássia Tibério Araújo e Maria Cláudia Cabrini Grácio da Faculdade de Filosofia e Ciências da Universidade Estadual Paulista, tecemos breves comentários.

É louvável a tentativa da condução de estudos comparativos avaliando o desempenho de estudantes que desenvolvem diferentes currículos, contudo devemos ter em mente que o olhar dirigido a projetos político-pedagógicos com ideologias e metodologias diversas pressupõe cuidados vitais na análise de quaisquer dados ou informações colhidas.

A visão disciplinar do currículo tradicional se contrapõe à óptica interdisciplinar do currículo em construção no Curso de Medicina da FAMEMA; a pedagogia da transmissão de conhecimentos contrasta-se com metodologias ativas como a aprendizagem baseada em problemas – ABP, onde não há "aprendizado", finito, estanque, mas sim aprendizagem, processo contínuo, busca permanente, que valoriza o estudante (de *studiu*, aplicação zelosa, ardor) enquanto sujeito de sua própria aprendizagem, em oposição ao "aluno" (de *alumni*, "criança que se dava para criar") que recebe passivamente a instrução.

Assim, submeter estudantes em processos e momentos distintos de formação a uma prova de conhecimentos com 40 testes de múltipla escolha, e comparar seu desempenho na disciplina de Nefrologia constitui de fato uma análise válida de "aproveitamento"?

Resposta a questões com "intensidade alta na categoria atualidade" permitem medir a utilização de periódicos pelos estudantes submetidos a exposições diferentes aos conteúdos avaliados?

Currículo e formação são movimento, e necessitam de estudos que registrem esta dinâmica. Eis o desafio da avaliação de estudantes que desenvolvem diferentes currículos: estamos querendo comparar o incomparável?

Ricardo Shoití Komatsu¹

Avaliação do Desempenho de Estudantes: uma Análise Crítica

Com a intenção de contribuir para o aprimoramento de pesquisas focalizando a verificação de aprendizagem de estudantes, trago alguns questionamentos para reflexão.

Os autores do artigo "Avaliação do aproveitamento do aluno em nefrologia: estudo comparativo entre o método expositivo tradicional e o método de ABP" traçaram como objetivo a identificação de diferenças no aproveitamento de estudantes de Medicina frente a 40 testes de múltipla escolha, sendo um grupo de 2ª série de um currículo em ABP – Aprendizagem Baseada em Problemas e outro de 4ª série, de um currículo tradicional, ambos da Faculdade de Medicina de Marília – Famema.

¹ Professor e Diretor de Graduação da FAMEMA – Faculdade de Medicina de Marília.

Chama a atenção que o número de testes utilizados ($n=40$) garante confiabilidade segundo Norman (1995). Porém, a classificação de cada categoria (profundidade, atualidade, capacidade de associação e conotação psicossocial), em 4 graus de intensidade, reduziu o número de questões nas classes, de forma a comprometer a confiabilidade (algumas classes ficaram com 2 questões).

Da mesma forma, 34 das 40 questões foram classificadas como tendo fraca ou nula conotação psicossocial. O fato de apenas 6 questões abordarem essa dimensão compromete, ou inviabiliza, a análise da terceira hipótese ("a abordagem psicossocial, determinada como enfoque obrigatório no método PBL, poderia constituir aspecto distintivo de desempenho dos grupos").

A classe com um maior número de questões foi a de profundidade com intensidade moderada ($n=33$). Porém, o que chama a atenção nessa categoria, que representa 82,5% do total do teste, é que as questões nela classificadas representam a "informação retida pelo aluno (...) em grau de complexidade comparável a de um especialista na área". Assim, cabe um questionamento em relação à validade do teste (Gronlund, 1971). Em relação à validade de conteúdo seria pertinente considerar o quanto a moderada intensidade na categoria profundidade corresponde aos objetivos educacionais estabelecidos na unidade educacional Sistema Renal e do Ciclo Nefro-Urologia, respectivamente para a 2ª e 4ª séries do curso médico da Famema e o quanto a retenção de informação, pelo estudante de Medicina, de forma comparável em grau de complexidade a de um especialista na área é um bom indicador de avaliação do aproveitamento, quer para a 2ª ou 4ª séries.

A inferência sobre o tipo de fonte utilizada pelos estudantes e a forma como eles constroem suas redes semânticas (conhecimento) precisaria de uma análise de hierarquia das questões utilizadas no teste, no tocante à taxonomia do domínio cognitivo. De acordo com a classificação de Bloom (1971) para esse domínio, poder-se-ia minimamente distribuir as questões do teste em dois grandes grupos: um fundamentado na memória e, portanto, na capacidade de lembrar fatos, eventos, princípios e mecanismos (como parece ser o caso das duas questões apresentadas como exemplo no artigo) e um outro grupo para questões de aplicação do conhecimento (compreensão, análise, síntese e avaliação).

Mesmo as questões categorizadas como de alta/moderada atualidade e/ou profundidade podem ser elaboradas de modo a avaliar a capacidade de memorização de informações ou de aplicação do conhecimento.

Assim, a segunda hipótese levantada pelos autores referente à suposição de que "alunos do PBL, ao serem estimulados à investigação, tendem a apresentar maior capacidade de aquisição de conhecimentos ou informações atualizadas, mesmo que estas não tenham sido fornecidas pelos professores especialistas", não esclarece o que significa aquisição de conhecimentos ou informações. O tipo de testes selecionados evidencia que tipo de diferenças os autores buscaram revelar/avaliar. Uma vez que não foram explicitadas o significado de aquisição de conhecimento ou a hierarquia predominante do domínio cognitivo abordada, a segunda hipótese não pode ser devidamente verificada ou testada.

Finalmente, em relação à primeira hipótese que focaliza os diferentes amadurecimentos e acúmulos de conhecimento entre estudantes da 2ª e 4ª séries, com vantagem para esses últimos, inúmeros estudos anteriores focalizando o desempenho progressivo dos estudantes por série expostos aos mesmos testes de múltipla escolha centrados em casos ou prática clínica confirma um percentual de acertos crescente. Porém, a segunda assertiva dessa hipótese não se correlaciona com a primeira, uma vez que o desempenho progressivo em testes de múltipla escolha já foi igualmente constatado para estudantes de currículos tradicionais como daqueles baseados em problemas. Apesar das questões acima colocadas, creio que esse é um tema de grande relevância e que permanece pouco explorado.

Valéria Vernaschi Lima²

REFERÊNCIAS BIBLIOGRÁFICAS

1. Gronlund NE. Measurement and evaluation in teaching. New York: The Macmillian Company, 1971.
2. Bloom B. (1971) Taxonomía de los objetivos de la educación. Buenos Aires: El Ateneo, 1971.
3. Norman G. Multiple Choice Questions. Evaluation methods: a resource handbook. Hamilton: Mc Master University, 1995.

² Professora e coordenadora do Comitê de Avaliação do Estudante - Faculdade de Medicina de Marília.