

AVALIAÇÃO DE COMPETÊNCIAS BÁSICAS

Itala Ferreira¹

Resumo — O autor relaciona estudos relativos a competências básicas no campo da educação médica, tanto no exterior, como no Brasil, para mostrar que existe consenso quanto à necessidade de definição de competências básicas para uma adequada avaliação do estudante de Medicina. Analisa instrumentos de medida apropriados para controle do processo de aprendizagem no que diz respeito ao domínio de uma competência — os denominados testes com referência a critério. Conclui, argumentando que a medida com referência a critério é meio para ajudar a resolver os problemas da medida do rendimento escolar e, como tal, permite constatar até que ponto o aluno se aproxima do domínio dos comportamentos necessários à sua profissionalização.

Introdução

São evidentes, na última década, as mudanças ocorridas na área da medida e da avaliação educacional. É de se notar a ênfase dada a estudos sobre competências básicas em todos os setores e níveis de ensino. Isso ocorre, conforme observa Berk⁸, "particularmente no campo da Medicina e demais profissões da saúde".

Pretende-se, pois, passar em revista alguns estudos sobre competências básicas realizados, tanto no exterior, quanto no Brasil, na área da Medicina.

Entre vários trabalhos significativos para o tema enumeram-se os dos seguintes autores: Hubbard¹⁹; Cowles¹⁴; Sanazaro³¹; Somolinos, Jinnich, Campillo, Valera y Sanches³⁰, preocupados em definir competências clínicas. Adler e Enlow², Barro⁶, Berner⁹, Aloia e Jonas³, que apresentaram estudos relacionados à competência diagnóstica; Randels²⁹, que estudou competências ligadas ao desempenho do psiquiatra; Semeonsson, Kenney, Walket³² e Burg¹¹, que apontaram competências ligadas à prática pediátrica; e Goetz, Peters, Folse, Beck, Fink e Wavak¹⁷, que definiram competências para identificar informações clínicas a partir de dados do prontuário.

O American Board of Internal Medicine⁴ descreveu as competências requeridas do clínico geral. Neste estudo, o clínico geral é conceituado como o médico que propicia, com bases científicas, cuidado empático para as doenças não cirúrgicas dos adultos. A explicitação das competências descritas é avaliada, através de exames rigorosos, com vistas à aquisição do certificado que confere o direito de clinicar. Categoriza as competências em atitudes, hábitos, habilidades e conhecimentos. O relatório não apresenta dados sobre a metodologia utilizada para o levantamento das competências, nem para a avaliação do candidato. Evidencia, entretanto, a necessidade de definição de competências básicas para seleção adequada dos profissionais que aspiram a clinicar.

Mc Gaghie²² propõe um modelo de plano de estudos de Medicina, baseado em competências, que se caracteriza por: (a) organização a partir de competências requeridas para exercício da Medicina em determinado ambiente; (b) fundamentação baseada no princípio, empiricamente comprovado, de que os estudantes podem alcançar o rendimento previsto desde que recebam ensino adequado; (c) concepção da educação como uma experimentação, no qual, tanto o processo de aprendizagem, como as estratégias utilizadas para sua eficiência, sejam consideradas como hipóteses sujeitas à comprovação e, portanto, objeto de pesquisa continuada.

O que se pretende obter de um programa baseado em competências é a formação de profissionais de saúde que possam exercer a Medicina, com certo nível de eficácia, e em consonância com as condições do ambiente, a fim de atender concretamente às necessidades locais.

Mc Gaghie²², em seguida, descreve: (a) os elementos que determinam a definição das competências; (b) os princípios que fundamentam a aprendizagem para o domínio de uma competência; (c) os tipos de avaliação da competência; e (d) a necessidade de preparar os docentes, os estudantes e as instituições de ensino médico para o desenvolvimento da nova estratégia.

(1) Coordenador, Programa de Preparo Pedagógico, ABEM.

No Brasil, alguns estudos sobre competências básicas na área médica começam a aparecer.

Bechara⁷ estabeleceu competências básicas em Protologia a serem demonstradas pelos médicos ao se graduarem na Faculdade de Medicina da Universidade Federal de Juiz de Fora. Este autor conceitua competências como, "afirmações descritivas de conhecimentos a serem demonstrados, por médicos generalistas, na execução de atividades relacionadas à sua profissão, definidas sob forma de objetivos gerais segundo a concepção de Gronlund". Identificou as principais doenças protológicas e os principais incidentes críticos, utilizando como fonte os prontuários do Hospital-Escola da Universidade e do INAMPS/JF. Em seguida, elaborou tabelas nas quais as competências são definidas sob forma de objetivos gerais e especificou situações em que as competências poderiam ser medidas. Posteriormente, as tabelas foram submetidas à apreciação de juízes. O pesquisador recomenda que "sempre que possível, os programas sejam propostos sob forma de competência".

Cabe, ainda, ressaltar o estudo realizado por um grupo de trabalho constituído pela Comissão Nacional de Residência Médica, da Secretaria da Educação Superior (SESu) do Ministério da Educação¹⁰, no qual é apresentada, sob o título de "Avaliação do Treinamento do Residente", uma proposta de instrumentos básicos para avaliação de residentes. Este trabalho classifica e descreve os instrumentos de medida, apresentando em primeiro lugar a "avaliação formativa por supervisão direta do desempenho e da competência". A avaliação é assim descrita: "trata-se de verificar, tanto o desempenho, como a competência, nos três níveis (cognitivo, psicomotor e afetivo), através da revisão diária dos casos (visitas, por exemplo) e demais atividades (consultas de pacientes externos, manuseio de aparelhos, colheita de material, prescrições etc.), sem registro escrito".

O relatório apresenta, ainda, outros instrumentos. Sugere vários modelos de fichas a serem preenchidas pelos preceptores e/ou médicos de instituição em que é feita a Residência, com a finalidade de avaliar os vários níveis de desempenho e competência. Deve ser enfatizado, entretanto, que o documento não permite perceber a metodologia empregada para a definição das competências esperadas do residente, nem relaciona as competências a serem avaliadas. As competências podem ser identificadas nas tarefas, atividades, atitudes, habilidades e nível de conhecimento discriminados nas fichas. Percebe-se, em todo trabalho, a preocupação com "a avaliação do desempenho global na produção de serviços de saúde" e a necessidade de definir as competências a serem demonstradas na Residência Médica.

Paula²⁴ analisa o processo de diagnóstico médico a partir do conceito de diagnóstico, identifi-

cando as habilidades que compõem a competência em diagnosticar e procurando definir os processos ou caminhos mais adequados para o diagnóstico. Define competência como: "conjunto de conhecimentos, de habilidades, de atitudes e também de esquemas cognitivos que são fundamentais para o exercício de uma atividade profissional e que estão virtualmente presentes no indivíduo competente, sendo atualizados no momento do desempenho".

O estudo visa a verificar por intermédio da Técnica de Delfos: (a) a evidência de consenso entre especialistas brasileiros sobre o diagnóstico médico; (b) a competência diagnóstica; (c) o processo de diagnosticar; (d) os objetivos e normas que devem reger o ensino do processo de diagnóstico.

Pode-se afirmar, no momento, que são numerosos os estudos já realizados nas universidades americanas a fim de definir competências básicas no ensino da Medicina.

Os estudos descritos revelaram, sem exceção, o consenso quanto à necessidade de definição de competências básicas para uma adequada avaliação do estudante de Medicina.

Conseqüentemente, justifica-se a construção de testes com referência a critério que têm como objetivo determinar o grau de alcance das competências previstas e definir o *status* de indivíduos, ou programas, em relação a padrões de desempenho pré-estabelecidos.

Medidas com referência a critério e competências básicas, segundo Popham

A fim de justificar o emprego dos testes com referência a critério para avaliação de competências básicas, há que conceituar e caracterizar, primeiramente, competências básicas segundo Popham, enfoque adotado neste estudo.

Popham²⁵, quanto à interpretação de competência básica, mostra que o ponto central de um programa com base em competências é a determinação das habilidades básicas que um estudante deverá demonstrar ao final de um curso.

Carvalho¹² comenta a respeito: "determinar quais devam ser essas habilidades é um problema crucial, pois, elas devem ser básicas o suficiente para serem exigidas de todos os alunos. Além disso, devem ser acompanhadas de níveis variados de desempenho, isto é, de um nível mínimo, esperado de todos os alunos, e de níveis elevados, esperados de alguns apenas".

Para Popham e Lindheim²⁸ uma competência básica descreve habilidade de natureza física, ou intelectual. Apontam, quanto ao conteúdo, quatro critérios a serem estabelecidos para o julgamento de uma competência: (a) mensurabilidade

de; (b) abrangência apropriada; (c) possibilidade de ser ensinada; (d) relevância.

A competência, pois, apresenta um comportamento que possa ser evidenciado, para ser medido. Determinadas habilidades, como, por exemplo, atitude de empatia face ao paciente, não podem ser consideradas competências. Habilidades do domínio afetivo, como o exemplo citado, serão avaliadas por outro tipo de instrumento.

Ter abrangência apropriada significa ser descrita de tal forma que possibilite a identificação de sub-habilidades e estas só poderão ser identificadas a partir de competência que não seja demasiado restrita.

Uma decorrência de característica acima descrita é a que se segue: a competência deve ser passível de ser ensinada, ou seja, deve ser de responsabilidade da escola. De acordo com Popham²⁵, um programa baseado em competências exige: (a) descrição das tarefas esperadas do estudante, quer dizer, o professor, desde o início de uma seqüência de ensino, descreve, para os estudantes, a habilidade terminal a ser atingida por eles. Supõe-se que, quando estes compreendem a natureza da habilidade a ser aprendida, tornam-se mais aptos para adquiri-la. Dessa forma, fornece-se ao aluno uma descrição da competência, em forma simplificada, no início da seqüência de ensino. Além disso, o professor despertará a motivação dos estudantes, mostrando-lhes a importância da habilidade. A compreensão clara, por parte do aluno, do que deles é esperado, possibilitará a concentração de seus esforços nas tarefas relevantes para o efetivo domínio das competências desejadas; (b) prática apropriada, isto é, ensinar o domínio de determinada competência específica, dar ao estudante muitas oportunidades de praticar o mesmo tipo de comportamento exigido para o domínio da competência; (c) conhecimento dos resultados. Da mesma maneira que se deve dar ao estudante oportunidade de praticar o comportamento a ser medido, o resultado do desempenho será conhecido, através de processos breves, para que tomem consciência de terem ou não atingido a habilidade prevista, quer intermediária, quer final; e (d) controle do processo, isto é, esquemas de controle continuado do progresso do aluno, objetivando criar oportunidades de recuperação paralela. Verificações de rendimento, com o objetivo de controle do progresso do estudante, devem ser empregadas com freqüência, de modo que, a partir da informação sobre a realidade do estudante, sejam tomadas as medidas apropriadas.

Finalmente, a competência deverá ser significativa, quer dizer, apresentar um valor intrínseco, julgado essencial ao indivíduo critério difícil de ser aplicado numa série de competências, quando se faz necessário estabelecê-las por ordem de prioridade.

Um programa de competências básicas exige instrumentos de medida apropriados para controle do processo da aprendizagem no que diz respeito ao domínio, ou não, de uma competência — os denominados testes com referência a critério. Este termo foi cunhado e introduzido por Glaser¹⁶ e reflete medidas em que o desempenho de um indivíduo é comparado a um critério, ou seja, a um domínio de conhecimentos pré-estabelecido. Para Millman²³, a medida com referência a critério tem duas características principais: (a) possibilita informações sobre "habilidades e conhecimentos específicos; e (b) fornece escores interpretáveis em termos de tarefas ou desempenhos". Popham²⁸, ao se referir à medida com referência a critério, a define como medida capaz de estimar o *status* do estudante em relação a um domínio de comportamento definido claramente.

Segundo Ebel¹⁵, testes com referência a critério são elaborados com a finalidade de medir o alcance de certo número de habilidades significativas e distintas entre si, a serem dominadas pelos estudantes, no decorrer de um programa. Dessa forma, para Ebel "o teste com referência a critério é, obviamente, o teste que deve ser usado num programa de competências básicas".

O esquema descritivo dos testes com referência a critério define com clareza a natureza do comportamento a ser medido, bem como permite a construção de questões homogêneas, ligadas a cada uma das competências preestabelecidas.

Informações indispensáveis para o programa de competências básicas, tais como a identificação do progresso do estudante, o diagnóstico das deficiências do processo e a avaliação das competências são proporcionados pelos testes com referência a critério, se construídos e interpretados adequadamente.

O rendimento do aluno num programa de competências básicas é medido em relação a um critério ou domínio de comportamentos. Os testes com referência a critério consideram o progresso individual do aluno, o ritmo próprio para a aprendizagem e, ainda, a importância da instrução, quando o estudante não atingiu objetivos intermediários ou sub-habilidades, relacionadas às competências básicas. Tornam, portanto, possível a individualização do processo ensino-aprendizagem.

Um programa de competências básicas tem seu principal motivo, entre todos os citados anteriormente, a interpretação objetiva dos resultados do aluno em cada competência pré-estabelecida, a partir de um padrão externo que orienta o julgamento do alcance, ou não, de tal habilidade, considerando cada estudante, não em comparação aos colegas, mas individualmente, em relação a "um domínio bem definido de comportamento"²⁵.

Construção de testes, com referência a critério

Popham^{2,5}, referindo-se ao planejamento de um teste com referência a critério, considera como seu atributo mais importante o esquema descritivo, também denominado especificação do teste. Quando elaborado com rigor, mostra com clareza o tipo de comportamento que o estudante será, ou não, capaz de evidenciar. Atribui à metodologia para especificação de domínios de conteúdo, duas funções principais: comunicar aos usuários a natureza do comportamento que está sendo medido e explicitar os pormenores do domínio do comportamento a ser medido, de modo que os construtores de teste possam gerar questões homogêneas, derivados da especificação.

Popham^{2,5} estabelece como componentes das especificações dos testes com referência a critério: (a) descrição geral; (b) amostra de questão; (c) atributos de estímulo; (d) atributos de resposta; e (e) especificações complementares.

A descrição geral tem como finalidade enunciar breve e precisamente o comportamento e a forma pela qual será medido. Fornece uma visão geral do comportamento a ser complementada mais detalhadamente nos componentes que se seguem. Como exemplo, apresenta-se a descrição geral utilizada na especificação do teste sobre a competência "conhecimento dos termos médicos utilizados na anamnese".

Descrição geral — Ao aluno serão apresentadas breves descrições que se referem a termos médicos. Após fazer a leitura de cada descrição, o aluno deverá identificar o termo médico correspondente à parte grifada da descrição apresentada.

A amostra de questão exemplifica a natureza das questões do teste, que são descritas pelos atributos de estímulo e atributos de respostas que compõem a especificação do teste. Inclui: (a) instrução para resposta; (b) estímulo ou enunciado da questão; e (c) alternativas de respostas. Exemplo:

Instruções: 1. Leia a descrição, atentamente.
2. Faça um "X" diante do termo médico utilizado para definir a parte sublinhada da descrição.

Ao exame, LFT, 5 anos de idade, apresentava *freqüência cardíaca de 40 batimentos por minuto.*

3. Qual o termo médico que corresponde à parte sublinhada da descrição?

- () Flictena
- () Bradicardia
- () Dispnéia paroxística
- () Taquicardia

Os atributos do estímulo, conforme Popham, são afirmações que, no seu conjunto, descrevem os fatores que norteiam a composição das questões, fixando os limites de seu conteúdo. Um exemplo:

1. Cada descrição do termo médico apresentado será precedido pelas seguintes instruções: 1. "Leia a descrição, atentamente". 2. "Faça um "X" diante do termo médico utilizado para definir a parte sublinhada da descrição". Essas instruções serão seguidas pelo estímulo: "Qual o termo médico que corresponde à parte sublinhada da descrição?"
2. Cada descrição terá de 8 a 50 palavras.
3. Os significados dos termos médicos apresentados nas descrições serão retirados de manuais de semiologia e dicionários de termos médicos. Serão feitas as devidas adaptações quando o número de palavras ultrapassar o limite estabelecido no item 2. Esta adaptação deve manter-se fiel à definição.
4. Os termos médicos cujas definições serão pedidas constam de listagem que faz parte da unidade Anamnese.

Atributos de resposta são constituídos por uma série de afirmações que consistem, conforme Popham, na "explicação dos critérios que permitam o julgamento fidedigno da adequação das respostas do examinado". Os atributos de resposta delimitam as classes de alternativas de onde o examinando seleciona as respostas ou explica os padrões para as respostas construídas. Em se tratando de respostas selecionadas, estas afirmações definem tanto as características das respostas incorretas, quanto as da correta. Exemplo:

1. O aluno receberá quatro alternativas de resposta.
2. A resposta certa, somente uma das opções, será a palavra que corresponder à definição do termo médico descrito e deve pertencer à unidade em estudo.
3. Os distratores poderão ser de dois tipos:
 - a) quando o termo médico selecionado para a resposta certa pertencer a um aparelho ou sistema específico, dois dos distratores serão relacionados ao mesmo aparelho ou sistema, e, o terceiro, a aparelho ou sistema diferente;
 - b) quando o termo médico selecionado para a opção certa for genérico, não se referindo a um aparelho ou sistema específico, os três distratores deverão ser da mesma natureza: genéricos.
4. A ordem em que a alternativa correta entra no rol de opções será escolhida aleatoriamente.

As especificações complementares constituem componente opcional, uma espécie de apêndice usado sempre que se torna necessário fornecer explicações mais detalhadas sobre o conteúdo a ser tratado. Uma especificação de teste é necessária para cada conjunto de questões que tenha por finalidade medir a mesma classe de comportamentos do examinando.

Abrahão¹ se posiciona em relação às especificações de testes: "Este esquema descritivo, além de orientar as questões de teste, permite a interpretação dos escores dos alunos em termos de comportamentos de entrada, pela aplicação do pré-teste; falhas ocorridas na instrução ou dificuldades do aluno pela aplicação do teste durante o processo de ensino e habilidades adquiridas, pela aplicação do teste após a instrução".

No planejamento e construção de testes com referência a critério, Popham^{2,7} recomenda a elaboração de número suficiente de questões em cada teste, de maneira que se torne possível a constatação do *status* do estudante em relação ao domínio do comportamento que se está medindo, sem contudo sobrecarregá-lo com excesso de questões. O número recomendado por Popham e Lindheim^{2,8} varia entre 10 e 20.

Validade dos testes com referência a critério

Considerando que o planejamento de um teste compreende também a sua validação e fidedignidade, uma vez construídas as questões do teste, torna-se necessário, antes de sua aplicação definitiva, usar estratégias que permitam, não só identificar as defeituosas, para correção, ou, se for o caso, para substituição, como também determinar o grau de consistência do instrumento.

Popham^{2,5} sugere três diferentes estratégias para a validação dos testes com referência a critério: (a) validade descritiva; (b) validade funcional; (c) validade de seleção do domínio.

Um teste com referência a critério apresenta validade descritiva quando as especificações estão de tal modo claras que possibilitam a geração de itens homogêneos. Popham^{2,5} condiciona a interpretação adequada do desempenho do aluno em um teste, à clareza de sua especificação.

Barbosa³ descreve como deve ser o procedimento para verificar a clareza das especificações: (a) solicita-se a duas ou mais pessoas que, a partir das especificações, elaborem itens. Se os itens elaborados apresentarem homogeneidade está comprovada a clareza das especificações; (b) solicita-se, numa segunda etapa, o julgamento de outro grupo de juízes quanto à congruência dos itens construídos. Os juízes deverão identificar, no conjunto de itens, os que apresentam homogeneidade. Alguns itens incongruentes serão intercalados propositalmente, entre os homogê-

neos, para se verificar a seriedade com que os juízes avaliam os itens apresentados.

A validade funcional está bem caracterizada por Popham^{2,6} quando diz: "entende-se por validade funcional a precisão com que o teste com referência a critério responde às finalidades para as quais foi elaborado". Exemplificando, um teste com referência a critério que tenha por objetivo medir habilidades para fazer uma anamnese completa, terá validade funcional na proporção em que os examinandos demonstrem ter atingido a habilidade tanto ao serem testados como posteriormente, ao longo da prática médica.

A validade de seleção do domínio ou validade de seleção do conteúdo é, conforme Popham^{2,7}, o tipo de validade que diz respeito à exatidão com que se seleciona determinado conteúdo com a finalidade de se estabelecer o *status* do examinando. O fundamental é que seja selecionado um domínio que, segundo Popham^{2,5}, represente, com exatidão, um indicador preciso das dimensões que o teste se propõe a medir".

Na prática, a validação de seleção do domínio se faz quando, definidas as competências, se solicita a um grupo de pessoas, conforme Barbosa⁵ "qualificadas no assunto em questão", que classifiquem, por ordem de importância, as competências.

Fidedignidade dos testes com referência a critério

Fidedignidade, conforme Ebel^{1,5}, consiste "no grau de consistência ou estabilidade de um instrumento que se propõe a medir". Sintetizando, fidedignidade é o grau de consistência de um instrumento de medida.

Conforme Carvalho¹²: "Testes com referência a critério bem construídos devem ser fidedignos, apresentando consistência de resultados em aplicações repetidas ou formas paralelas do teste".

Hambleton, Swaminathan, Algina e Coulson¹⁸ apresentam, em relação à fidedignidade de um teste com referência a critério, três possibilidades de investigá-la: (a) por consistência de classificação em posição de domínio ou não domínio; (b) por consistência dos escores do teste; e (c) por consistência dos escores individuais.

O primeiro procedimento diz respeito à consistência com que os alunos são classificados em posição de domínio, ou não, em mais de uma aplicação do teste. Estabelecido um padrão de desempenho, também denominado *escore de corte*, ponto de corte ou *cut-off*, se os alunos se posicionarem consistentemente acima ou abaixo do padrão determinado, em duas aplicações consecutivas ou simultâneas de formas paralelas do teste, pode-se concluir que o teste é fidedigno.

O segundo enfoque também exige duas aplicações consecutivas do mesmo teste ou simultânea de duas formas paralelas, e se detém sobre a verificação dos desvios dos escores dos alunos em relação ao padrão de desempenho.

O terceiro enfoque descreve a consistência dos escores individuais, também por duas aplicações do mesmo teste ou de formas paralelas sem, entretanto, considerar o padrão de desempenho.

Das três abordagens acima descritas, a mais adequada ao teste com referência a critério é a que diz respeito à consistência com que os estudantes são classificados em posições de domínio ou não.

Klein²¹, visando à consistência de classificação em posição de domínio e não domínio, utilizou três estimadores: (a) o estimador sugerido por Hambleton e Novick, citado por Klein; (b) o estimador determinado por Swaminathan, Hambleton e Algina³³; e (c) o estimador proposto por Huynh²⁰.

Hambleton e Novick, citados por Klein²¹ sugerem duas aplicações consecutivas do teste, para se obter o índice de fidedignidade de consistência de classificação em posição de mestria ou não. O Índice de fidedignidade, de acordo com Klein é obtido através de fórmula P_0 que significa proporção observada de estudantes classificados na posição de mestria ou não.

Swaminathan, Hambleton e Algina³³ argumentam que além da fórmula P_0 , deve-se levar em conta a proporção de decisões consistentes ocorridas por chance. E sugerem o uso do coeficiente K (Cohen¹³).

Huynh²⁰ propõe um índice de fidedignidade obtido através da fórmula K_1 , que exige uma única aplicação do teste.

Trata-se de um estimador que pode ser considerado como adaptação do índice K. Seu modelo é adequado particularmente quando aplicado a estudantes ensinados para a aquisição de habilidade.

Deve-se ressaltar que a medida com referência a critério, por si só, não resolve todos os problemas da medida do rendimento escolar. Entretanto, preparada e aplicada adequadamente, não apenas permite a definição de quanto o aluno se aproxima do domínio de comportamento para os quais o ensino se volta, quanto apresenta implicações para o próprio processo de ensino.

Summary — *The author discusses several studies on basic competencies in the field of medical education, in Brazil as well as in other countries, which show a consensus about the need of defining basic competencies for an adequate evaluation of the medical student. She analyzes appropriate measurement instruments for monitoring the learning process in relation to one single competence — the so-called criteria-related tests. She concludes that criteria-related measurements are only means to help solve the problem of measurement of performance at the school, and as such, to permit to verify to what extent*

the student has reached those behaviors needed for his performance as a professional.

Referências Bibliográficas

1. ABRAHÃO, Maria Christina. *Competências básicas em leitura ao final da 4.ª série do 1.º grau: O básico alcança padrões mínimos?* (Dissertação de Mestrado). Rio de Janeiro: Faculdade de Educação, Universidade Federal do Rio de Janeiro, 1980.
2. ADLER, L. M. & ENOLW, A. J. An instrument to measure skill in diagnostic interviewing: a teaching an evaluation tool. *Journal of Medical Education*, 41, 1152-7, 1966.
3. ALOIA, John F. & JONAS, Ernesto. Skills in history — talking and physical examination. *Journal of Medical Education*. 51, 410-15, 1976.
4. AMERICAN BOARD MEDICINE: Philadelphia, Pennsylvania. Clinical competence in internal medicine. *Annals of Internal Medicine*, 90, 402-11, 1979.
5. BARBOSA, Angela Maria Magnan. *Construção e validação de testes com referência a critério para a medida de competências relacionadas à língua portuguesa, a serem demonstradas por alunos do 1.º grau.* (Dissertação de Mestrado). Rio de Janeiro: Faculdade de Educação, Universidade Federal do Rio de Janeiro, 1980.
6. BARRO, Arlene R. Survey & evaluation of approaches to physician performance measurement. *Journal of Medical Education*, 48, 1053-93, 1973.
7. BECHARA, Alfredo Salomão. *Competências básicas em protologia a serem demonstradas pelo médico generalista, ao graduar-se na Faculdade de Medicina da Universidade Federal de Juiz de Fora.* (Dissertação de Mestrado). Rio de Janeiro: Faculdade de Educação, Universidade Federal do Rio de Janeiro, 1979.
8. BERK, Ronald A. Item Introduction. In R. A. Berk (Ed). *Criterion — referenced measurement: The state of the art.* Baltimore: The Johns Hopkins University Press, 1980.
9. BERNER, E. S. Relationship between verbalization of problem-solving and problem-solving examination scores. *Journal of Medical Education*, 51, 1023, 1976.
10. BRASIL, Ministério da Educação e Cultura. Secretaria de Ensino Superior. Comissão Nacional de Residência Médica. Conclusões dos grupos de trabalho para elaboração de subsídios aos programas de residência médica. *Residência Médica*, 1, 80-97, 1980.
11. BURG, F. D. et alii. A method for defining competency in pediatrics. *Journal of Medical Education*, 51, 824-8, 1976.

12. CARVALHO, Maria Vicentina de Brito. *Aprendizagem da matemática ao final da 4.ª série do 1.º grau: Rendimento mínimo em competências básicas*. (Dissertação de Mestrado). Universidade Federal do Rio de Janeiro, 1981.
13. COHEN, I. A. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1, 37-46, 1980.
14. COWLES, A. A critical-comments approach to the rating of medical students clinical performance. *Journal of Medical Education*, 40, 188-98, 1965.
15. EBEL, Robert L. *Essentials of education measurement*. Englewood Cliffs, New Jersey: Prentice-Hall: 1972.
16. GLASER, Robert. Instructional technology and the measure of learning outcomes. *American Psychologist*, 18, 519-21, 1963.
17. GOETZ, A. et alii. Chart review skills: A dimension of clinical competence. *Journal of Medical Education*, 54, 788-96, 1979.
18. HAMBLETON, R. K., et alii. Criterion-referenced testing and measurement: A review of technical issues and developments. *Review of Educational Research*, 1, 1-47, 1978.
19. HUBBARD, J. P. et alii. An objective evaluation of clinical competence. *The New England Journal of Medicine*, 272, 1321-8, 1965.
20. HUYNH, Huynh. On the reliability of decisions in domain — referenced testing. *Journal of Educational Measurement*, 13, 254-64, 1976.
21. KLEIN, Maria Helena C. *Estimadores de fidedignidade em testes com referência a critério: Comparação quanto a pressupostos estatísticos e coeficientes obtidos em aplicação a dados hipotéticos*. (Dissertação de Mestrado). Rio de Janeiro: Faculdade de Educação, Universidade Federal do Rio de Janeiro, 1979.
22. MC GAGHIE, William C., et alii. *Introducción a la preparación de planes de estudios de medicina basados en la competencia*. Ginebra: Organización Mundial de la Salud, 1978.
23. MILLMAN, Jason. Criterion-referenced measurement. In William J. Popham (Ed). *Evaluation in education: Current applications*. Berkeley, California: McCutchan, 1974.
24. PAULA, Marcelo Alves de. *Estudo sobre a competência diagnóstica — uma contribuição do ensino e aprendizagem do diagnóstico na graduação médica*. (Dissertação de Mestrado). Rio de Janeiro: Pontifícia Universidade Católica, 1980.
25. POPHAM, William J. *Criterion referenced measurement*. Englewood Cliffs, New Jersey: Prentice-Hall, 1978a.
26. POPHAM, William J. *Educational evaluation*. Englewood Cliffs, New Jersey: Prentice-Hall, 1975.
27. POPHAM, William J. Well-crafted criterion — referenced tests. *Educational Leadership*, 1978b, 2, 91-5.
28. POPHAM, William J. & LINDHEIM, Elaine. *Making minimum competency programs work*. Los Angeles: Instructional Objectives Exchange, 1979.
29. RANDELS, P. M. et alii. Comparison of the psychiatry learning system and traditional teaching of Psychiatry. *Journal of Medical Education*, 51, 751-7, 1976.
30. SOMOLINOS, et alii. El expediente clínico y su importancia en la asistencia y educación médicas. *Gaceta Médica de Mexico*, 11, 477-91, 1980.
31. SANAZARO, P. J. & WILLIANSO, J. W. A classification of physician performance in internal Medicine. *Journal of Medical Education*, 43, 389-97, 1968.
32. SIMEONSSON, Rune, J., KENNEY, William and WALKER, Linda. Child development and disability: Competency — based clerkship. *Journal of Medical Education*. 51, 578-81, 1976.
33. SWAMINATHAN, H. HAMBLETON, R. K. & ALGINA, J. Reliability of criterion referenced tests: A decision — theoretic formulation. *Journal of Educational Measurement*. 11, 263-7, 1974.

Endereço do autor:

R. Afonso Cavalcanti, 13/201
20211 — Cidade Nova
Rio de Janeiro — RJ