

# Confiabilidade dos julgamentos de avaliadores de prova escrita na seleção para o mestrado

---

TAHÍSE MARIA DE BRITO MEDEIROS DOS SANTOS\*  
VIVIANE RODRIGUES VIANA MONTEIRO\*\*  
JOSÉ FLORÊNCIO RODRIGUES JUNIOR\*\*\*

## RESUMO

Setores da comunidade acadêmica têm expressado seu entendimento de que programas de pós-graduação, particularmente os de *stricto sensu*, devem-se submeter a autoavaliações. O presente estudo insere-se num contexto de autoavaliação de um programa de mestrado. Seu propósito é investigar o perfil de confiabilidade dos julgamentos dos examinadores da prova escrita no processo de seleção ao Mestrado em Educação de uma universidade privada do Centro-Oeste. O intervalo considerado abrangeu os anos de 1994 a 2006, totalizando 10 processos de seleção. Analisaram-se os dados numéricos das atribuições de notas, ano a ano, com base no coeficiente de correlação de Pearson. Os resultados apresentam um perfil irregular, com índices de correlação tanto baixos ( $r=0,15$ ) como elevados ( $r=0,89$ ). A análise estritamente documental mostrou-se insuficiente para explicar o padrão errático das correlações ao longo dos dez processos examinados. Sugerem-se procedimentos adicionais, tanto no que se refere ao pesquisar como no que diz respeito à elevação da confiabilidade no aferimento das provas escritas do processo de seleção.

**Palavras-chave:** Mestrado, Pós-graduação, Autoavaliação, Avaliação de desempenho.

---

\* Mestre em Educação, Universidade Católica de Brasília (UCB) (tahise\_britomedeiros@yahoo.com.br).

\*\* Aluna do Mestrado em Educação, Universidade Católica de Brasília (UCB) (vr.viana@uol.com.br).

\*\*\* Ph.D., Universidade de Brasília (UnB) (jffloren@apis.com.br).

## RESUMEN

Algunos sectores de la comunidad académica se han mostrado favorables a someter los programas de postgrado, particularmente los de *stricto sensu*, a autoevaluaciones. El presente estudio se inserta en un contexto de autoevaluación de un programa de máster. Su propósito es investigar el perfil de confiabilidad de las decisiones de los examinadores de la prueba escrita en el proceso de selección al Máster en Educación de una universidad privada de la región Centro-Oeste de Brasil. El intervalo considerado abarcó de 1994 a 2006 y se ocupó de 10 procesos de selección. Se analizaron los datos numéricos de las notas atribuidas año a año en base al coeficiente de correlación de Pearson. Los resultados presentaron un perfil irregular, con índices de correlación tanto bajos ( $r=0,15$ ) como elevados ( $r=0,89$ ). El análisis estrictamente documental se mostró insuficiente para explicar el estándar errático de las correlaciones en los diez procesos examinados. Se sugieren procedimientos adicionales, tanto en lo que se refiere a investigar como con respecto al aumento de confiabilidad en la apreciación de las pruebas escritas del proceso selectivo.

**Palabras clave:** Máster, Posgrado, Autoevaluación, Evaluación del desempeño.

## ABSTRACT

Representatives of the academic community have expressed their understanding that graduate programs should submit themselves to self-evaluations. This study is part of the self-evaluation scenario previously referred to. It intends to investigate how reliable the examiners' judgments are of a written exam, one of the components of the selection process to a Master of Education program of a private Midwest Brazilian university. The time span considered in the study was 1994 to 2006, comprising 10 entrance exams. The data were submitted to a Pearson correlation test year by year. The results showed an irregular pattern, in that some of the coefficients were very low ( $r=0,15$ ), some very high ( $r=0,89$ ). An analysis of solely the academic records proved insufficient to explain the erratic pattern of the correlations found in the 10 selection processes. Suggestions are presented concerning, on one hand, the research approach; on the other hand, ways of increasing the reliability of the grading process of the written exam making up the selection process.

**Keywords:** Master's Degree, Post-graduation, Self-evaluation, Performance evaluation.

## INTRODUÇÃO

Avaliação pode ser definida com um processo de delinear, obter e prover informação que permita julgar alternativas de decisão (Stufflebeam, 1971). Uma avaliação deve, entre outros objetivos, orientar as ações no setor a que se refere, determinar se as atividades do programa atendem aos objetivos declarados, se são apropriadas e se estão sendo efetivamente implementadas. Tomando-se essa perspectiva como ponto de partida, a avaliação de um programa pode permitir que os responsáveis pelas decisões meçam sistematicamente suas atividades — neste caso, o processo de seleção da pós-graduação. Avaliar é, portanto, um processo tão importante quanto complexo, e demanda permanente escrutínio para aperfeiçoá-lo e torná-lo uma ferramenta eficaz.

A importância dos processos de avaliação é considerada estratégica como subsídio indispensável no monitoramento das reformas e das políticas educacionais. Qualquer país preocupado em aumentar a eficiência, a equidade e a qualidade do seu sistema educacional necessitará investir em mecanismos de acompanhamento dos processos de reforma.

O papel da avaliação como motor para a melhoria de processos e instituições educacionais nos cursos de pós-graduação no Brasil vem sendo estudado por pesquisadores ligados à Coordenação de Aperfeiçoamento do Pessoal do Ensino Superior (Capes). Em particular, a autoavaliação constitui um dos critérios estabelecidos pelo referido órgão.

Duas pesquisas evidenciam a importância da autoavaliação como marca de programas que buscam a elevação de sua qualidade. Spagnolo e Souza (2004) efetuaram um estudo sobre as possibilidades de mudanças nos critérios da avaliação dos cursos *stricto sensu* brasileiros pela Capes. Foram consultados os membros das comissões avaliadoras e os coordenadores de programas de pós-graduação de instituições privadas. O estudo revelou que os dois grupos de pesquisadores concordam fortemente que se deve “utilizar a autoavaliação para a promoção da qualidade” (Spagnolo; Souza, 2004). A segunda pesquisa, realizada com o colegiado de doutores da Fundação Oswaldo Cruz, tomou como base uma série de proposições sugeridas por observadores internacionais convidados pela Capes para analisar a última avaliação da pós-graduação no Brasil em 2001 (Hortale, 2003). O estudo revelou a existência escassa de unidades internas de autoavaliação nas IES brasileiras. Além disso, quando essas existem, as informações por elas originadas não refluem para os processos das mesmas instituições.

A mesma ênfase no monitoramento com vistas à eficiência das instituições sociais é posta por De Bem (2004). Em um estudo de caso sobre a confiabilidade e

validade estatística da avaliação docente pelo discente, o autor afirma: “Diversas são as definições de avaliação que mencionam que as informações decorrentes do processo avaliativo devam apresentar as características de confiabilidade e de validade” (2004, p. 148). Como exemplo, cita a Conferência Mundial sobre o Ensino Superior, que define a avaliação como um “processo de coleta e tratamento de informações pertinentes, válidas e confiáveis, permitindo aos agentes responsáveis tomar as decisões que se impõem, para melhorar as ações e os resultados” (apud De Bem, 2004, p. 27). Propõe ainda, o referido autor, que os responsáveis pelo setor de avaliação das instituições educacionais investiguem a qualidade de seus dados, utilizando procedimentos empíricos e oferecendo, desse modo, maior segurança aos responsáveis pela tomada de decisão.

Para cumprirem seus propósitos as avaliações precisam atender a requisitos. A literatura sobre avaliação educacional indica critérios para esse propósito. Por exemplo, Mehrens e Lehmann (1978), adicionam ao critério **validade**, o critério **confiabilidade** na constituição de uma avaliação de boa qualidade. Como se recorda, validade implica aferir aquilo que efetivamente se quer medir. Por outro lado, confiabilidade pode ser entendida como a capacidade de discriminar, ou seja, a capacidade que um instrumento tem de separar indivíduos ou objetos que possuem um determinado traço ou propriedade de outros que não o têm. Ela pode ser ainda entendida como consistência interna da avaliação.

## **BOSQUEJO SOBRE A SELEÇÃO AO MESTRADO EM EDUCAÇÃO**

O processo de seleção de mestrado no Brasil é norteado pelo parecer n. 977/65 do Conselho Federal de Educação, que em seu art. 10 preconiza: “Para matrícula nos cursos de pós-graduação, além do diploma de graduação, as instituições deverão estabelecer requisitos que assegurem rigorosa seleção intelectual dos candidatos”. Estudos acerca do processo de seleção para o Mestrado em Educação em universidades brasileiras indicam a adoção de critérios variados, tais como: projetos de pesquisa, prova de língua estrangeira, análise do curriculum vitae, entrevista e prova projeto (Rodrigues Junior; Carnielli; Moura, 2006).

A partir da Lei n. 9.394/96, art.92, que exige que as Instituições de Ensino Superior (IES) apresentem, em seus quadros, profissionais docentes com titulação mínima de mestrado, a demanda por ingresso em programas de mestrado vem-se tornando crescente. No período de 1996 a 2003, o ingresso nos mestrados, no território nacional, elevou-se de 45.622 para 72.001, representando um aumento da ordem de 63,36% (Brasil, 2004).

A instituição cujo Mestrado em Educação é objeto deste estudo, teve seu primeiro curso de pós-graduação *stricto sensu* iniciado em 1994. Ao ser criado, objetivou a formação do docente da educação superior, do pesquisador e do administrador em educação. Além do docente do seu próprio quadro, necessitando adquirir a titulação mínima exigida, a instituição procurou, por meio do mestrado, proporcionar ao profissional em educação do Centro-Oeste a oportunidade de melhor qualificar-se.

## PESQUISAS SOBRE SELEÇÃO AO MESTRADO EM EDUCAÇÃO

Revisam-se nesta seção três pesquisas sobre o processo de seleção ao Mestrado em Educação no país. Professores do Mestrado em Educação, professores da graduação e alunos aprovados na seleção para o Mestrado em Educação de uma universidade pública do Centro-Oeste integraram a amostra da pesquisa de Ordeix (1982). A autora confrontou respostas dadas pelos três grupos quanto aos aspectos **utilidade, operacionalização e caráter eliminatório-classificatório** de nove critérios de avaliação praticados na seleção. No julgamento dos três grupos com respeito à **utilidade** dos nove critérios verificou-se convergência entre os grupos sobre os critérios indispensáveis: **formação profissional, pensamento analítico e crítico, motivação e disponibilidade de tempo**.

Ordeix (1982) constatou também a validade preditiva dos critérios de seleção dos critérios indicados como indispensáveis com relação às variáveis dependentes **desempenho geral e desempenho cognitivo**. Os coeficientes de correlação encontrados pela pesquisadora satisfizeram o ponto de corte de probabilidade igual ou superior a cinco por cento apenas para o critério **formação profissional**. Esse achado de Ordeix realça a importância do preparo profissional do candidato ao Mestrado em Educação.

O estudo de Matos (1995) pesquisou, por meio de questionários, os objetivos e critérios para a seleção ao Mestrado em Educação e Psicologia, no entendimento de alunos e professores, em três universidades de São Paulo. Da análise dos dados, identificaram-se os seguintes critérios: (a) **noção de metodologia**, (b) **experiência com pesquisa**, (c) **atitude científica**, (d) **boa base de informação**, (e) **domínio de línguas**, (f) **intenção de pesquisa executável**, (g) **boa formação na graduação**, (h) **criticidade, criatividade e persistência**, (i) **capacidade para elaborar e executar projetos de pesquisa**, (j) **habilidade para escrever com clareza e argumentar com lógica e coerência**, (k) **leitura com entendimento de texto técnico**, (l) **competência em utilizar bibliotecas e fazer levantamento bibliográfico**, (m) **compatibilização com a área de concentração e interesse pela pesquisa**, (n)

**atuação na docência do 3º grau, (o) disponibilidade de tempo, (p) currículo, (q) bom desempenho nos cursos introdutórios, (r) potencial para transferência do conhecimento e (s) interesse social.** Como se vê, os alunos e professores pesquisados por Matos (1995) indicaram 19 critérios de seleção para Mestrados em Educação e Psicologia. Entretanto, o pesquisador não logrou ordenar os referidos critérios, limitando-se a identificá-los.

O estudo de Carnielli e Rodrigues Júnior (2004) destinou-se a identificar instrumentos e procedimentos de seleção ao Mestrado em Educação no Brasil. Esse estudo foi realizado a partir dos editais de 44% dos programas de Mestrado em Educação, registrados à época na Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e disponíveis na internet. Obtiveram como resultados a existência de um núcleo de procedimentos onipresentes: **prova escrita, entrevista e *curriculum vitae*.**

É preciso destacar que os instrumentos utilizados no processo de seleção ao mestrado, particularmente o Mestrado em Educação, por mais bem planejados que sejam nunca são perfeitos e, desse modo, os dados coletados são suscetíveis de imprecisão, que se reflete nos julgamentos. Acrescentam-se a essa circunstância limitações, tais como: ausência de instrumentos padronizados, falta de normas, pressão do tempo, dúvidas quanto à confiabilidade dos julgamentos, e outras limitações as quais podem contribuir para que as informações levantadas nem sempre sejam inteiramente adequadas ao processo decisório de identificação dos mais capazes.

## **O CONTEXTO DA PRESENTE PESQUISA**

Visando ao aprofundamento na investigação sobre o processo de seleção, a presente pesquisa levanta a seguinte questão: **Diferentes juízes no processo de avaliação, julgando o mesmo componente de seleção – por exemplo, a prova-projeto –, têm julgamentos próximos?**

Um dos principais critérios para a avaliação da qualidade de um instrumento de mensuração é a sua confiabilidade e esta pode ser avaliada pelo grau de coerência com a qual esse instrumento mede o atributo (Polit; Hungler, 1995). A confiabilidade será considerada maior quando, em repetidas mensurações de um atributo, se obtiver uma baixa variação. Pode-se também dizer, que um instrumento é confiável considerando-se o seu grau de precisão, ou seja, quando suas medidas conseguem refletir, de forma precisa, as medidas reais do atributo investigado. A determinação da confiabilidade fornece uma indicação da extensão na qual a medida de um instrumento admite erros variáveis em seu uso (Horn, 1986; Polit; Hungler, 1995).

O conceito de confiabilidade utilizado, no presente estudo, refere-se à congruência de julgamentos de diferentes avaliadores, ou, ainda, à confiabilidade interavaliadores (*interrater reliability*). Como lembram Quellmaltz (1990) e Perroca e Gaidzinski (2002), esse índice possibilita a verificação do grau de correspondência entre as avaliações independentes de dois ou mais avaliadores. O presente estudo pretende verificar o grau de correspondência entre pares de juízes ao atribuírem notas à prova escrita. Em linguagem estatística, trata-se, portanto, de um procedimento de **correlação**.

Um estudo sobre a confiabilidade interavaliadores (*interrater reliability*), em entrevistas altamente estruturadas, aponta correlações moderadas (0,67); este coeficiente baixa para 0,34 quando as entrevistas não são estruturadas (Conway, Jako; Goodman, 1995). Diferentemente, pesquisa de Perroca e Gaidzinski (2002), no intuito de estudar, analisar e testar a confiabilidade do instrumento de classificação de pacientes por enfermeiras em hospital-escola, verificou correlações altas e positivas entre as avaliadoras; nesta pesquisa utilizaram-se listas de verificação (*check-list*) para classificar os pacientes.

## **METODOLOGIA**

Com o objetivo de aferir o grau de confiabilidade dos julgamentos dos avaliadores da prova escrita, utilizou-se nesta investigação a pesquisa documental para levantamento dos dados, conjugada com a análise estatística. Por meio da atribuição de números a fenômenos sociológicos, políticos, enfim, fenômenos ligados às Ciências Sociais, é possível mensurar os referidos fenômenos e, subsequentemente submetê-los a tratamento estatístico. A estatística, por sua vez, permite testar as relações dos fenômenos entre si e obter generalizações sobre sua natureza, ocorrência ou significado (Lakatos; Marconi, 2005).

Para verificar a existência seja de convergência, seja de discrepância entre os julgamentos dos avaliadores em linguagem estatística, utilizou-se o procedimento de correlação de Pearson. Esse procedimento estatístico constitui-se no método mais comum para descrever a associação entre variáveis quantitativas. O coeficiente de correlação designa a magnitude de relação entre duas variáveis, mensuradas em escala intervalar; é preciso observar que o tratamento intervalar dado ao resultado das provas escritas representa um artifício. Isto porque não se pode assumir que, por exemplo, a nota 8,0, atribuída a uma prova, seja equivalente ao dobro de uma nota 4,0, atribuída a outra prova.

Como é sabido, os coeficientes de correlação são, normalmente, designados com um  $r$  e costumam variar de +1,00 (para uma associação direta absoluta entre duas

variáveis) passando por 0,00 (ausência de relação) e chegando a -1,00 (para uma associação inversa absoluta entre duas variáveis). Dessa forma, quanto mais elevado for o valor obtido, mais confiável pode ser considerado o instrumento (Polit; Hungler, 1995). Para análise dos resultados, utilizaram-se os parâmetros descritos por Bauer e Gaskell (2002). Com base neles, os valores são considerados muito altos quando  $r > 0,90$ , alto quando  $r > 0,80$  e aceitável na amplitude  $0,66 < r < 0,79$ .

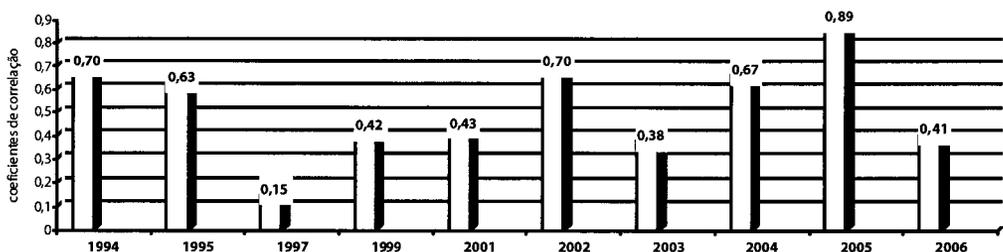
## COLETA DE INFORMAÇÕES

As notas da prova escrita foram obtidas com base na documentação oficial fornecida pela secretaria da pós-graduação da IES. As pastas contendo a documentação referente aos processos de seleção do Mestrado em Educação estavam disponíveis e separadas de acordo com o ano da seleção (1994, 1995, 1997, 1999, 2001, 2002, 2003, 2004, 2005 e 2006). Cabem duas observações a respeito dos registros examinados. Primeira, o julgamento da prova escrita, na maioria dos processos, foi realizado por dois examinadores e apenas em 1994 e 1995 foi feito por três examinadores. Segunda, observou-se que os registros do processo de seleção do período de 1994 a 2001 estavam manuscritos. Diferentemente, os dados correspondentes aos anos de 2002 a 2006 estavam digitalizados, facilitando ao pesquisador, tanto o acesso a eles quanto a sua análise. Depois da leitura dos documentos, procedeu-se à coleta, digitação e análise dos dados, utilizando-se a planilha Excel, assim como o programa *Statistical Package for the Social Sciences* (SPSS), versão 11.0.

## RESULTADOS E DISCUSSÃO

Os dados serão apresentados a seguir, considerando-se o ano do processo de seleção e o coeficiente de correlação obtido entre os juízes na correção da prova escrita. O Gráfico 1 mostra o coeficiente de correlação entre os avaliadores nos respectivos anos.

**Gráfico 1** – Correlações entre os avaliadores da prova escrita por ano



A análise dos índices de correlação de Pearson, tomando-se como base a classificação proposta por Bauer e Gaskell (2002), revelam correlação alta entre as notas atribuídas pelos pares de avaliadores em 2005; correlações aceitáveis em 1994, 1995, 2002 e 2004; e correlações baixas em 1997, 1999, 2001, 2003, 2006.

Considerando-se os resultados obtidos com base nos coeficientes de correlação entre os pares de avaliadores da prova-projeto, de maneira sucinta, pode-se destacar os seguintes aspectos:

- Nos dois primeiros anos do processo de seleção, a confiabilidade dos julgamentos foi considerada aceitável, comparativamente aos processos de 1997, 1999 e 2001. É importante levar em conta que nos dois primeiros anos de seleção, a prova escrita foi avaliada por três juízes, além da adoção do *tertius*. O processo de seleção incluiu a realização de seminários com presença obrigatória dos selecionados para posterior aplicação da prova escrita. Esse dado pode justificar, em parte, uma maior coesão teórica entre os selecionados e consequente proximidade na avaliação dos juízes.
- Em 2002, a comissão responsável pela seleção decidiu pautar o procedimento de aferimento das provas tomando como base indicadores objetivos; esses indicadores foram consolidados numa ficha de instruções. É possível que essa circunstância explique a melhora no índice de correlação entre os avaliadores naquele ano.
- Em 2003 e 2006 houve uma baixa nas correlações entre as notas atribuídas pelos avaliadores (Figura 1). Considerando, primeiro, a similaridade entre os processos de seleção e, segundo, que o processo seguiu com rigor o preceituado nos editais, é aparente a necessidade de um aprimoramento avaliativo dos juízes. Em contrapartida, a confiabilidade entre os avaliadores no processo de seleção de 2005, foi considerada muito alta (0,89). A irregularidade nos índices de correlação sugere a necessidade de os avaliadores, tendo em vista a existência de critérios objetivos, tal como configurado na ficha de avaliação, fazer uso dos mesmos com rigor.

Ressalte-se o fato de existirem no estudo limitações que impediram uma maior elucidação do problema levantado. Apontam-se duas. A coleta dos dados restringiu-se aos registros documentais. Durante a análise, verificou-se a existência de situações não explicáveis com base apenas naqueles documentos. Tornou-se evidente aos pesquisadores a necessidade de proceder-se também a uma investigação

histórica. Ou seja, é preciso ouvir os participantes do processo de seleção, particularmente nos anos iniciais. Questionamentos suscitados no curso da análise não foram respondidos pela ausência de registros mais acurados e pela impossibilidade de localizar as pessoas já referidas.

A segunda limitação em aferir a confiabilidade entre juízes da prova escrita decorre da subjetividade dos mesmos. Explicitando, cada docente avaliador tem sua história, seus valores, formação cultural e acadêmica idiossincráticos. Sendo pessoas diferentes, não se pode esperar que façam julgamentos absolutamente convergentes dos textos a eles submetidos.

Uma sugestão a ser considerada refere-se à prática usual, nos processos de seleção, de submeter a prova avaliada com resultados discrepantes a um terceiro juiz (*tertius*). Como visto no processo, tal atitude favorece a confiabilidade no julgamento das provas e, portanto, do processo de seleção.

No início desta comunicação procurou-se realçar a importância da autoavaliação dos programas de mestrado como instrumento capaz de promover a elevação de sua qualidade. Esse desiderato tem sido reiterado pela Coordenação de Aperfeiçoamento de Pessoal do Ensino Superior (Capes) em documentos recentes. Entretanto, como observadores e pesquisadores atestam, a autoavaliação de programas de mestrado tem existência formal, particularmente nas universidades federais; porém, quando se realiza, seus resultados não refluem para incremento e aperfeiçoamento dos programas. Como assinala Saorim (2009), citando entrevistados de sua pesquisa, “a autoavaliação é imprescindível, é fundamental... [porém] quando não leva em consideração a autoavaliação tira o poder dos envolvidos diretamente na formação constatarem os problemas e as possibilidades de seus cursos, as diferenças de áreas, cursos e região” (p. 119). Assim, nesta comunicação, objetivou-se chamar a atenção para a necessidade da autoavaliação em programas de mestrado, particularmente no tocante aos seus processos de seleção.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BAUER, W. M.; GASKELL, G. *Pesquisa qualitativa com texto, imagem e som: um manual prático*. 5.ed. Petrópolis: Vozes, 2002.
- BRASIL, Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. *Plano Nacional de Pós-Graduação (PNPG)*, 2005-2010. Brasília: MEC/Capes, 2004.
- CARNIELLI, B. L.; RODRIGUES JÚNIOR, J. F. *Perfil do candidato ao mestrado em educação: tendências e percepções*. In: SEMANA UNIVERSITÁRIA, 6., 2004. Brasília: UCB, 2004.

- CONWAY, J. M.; JAKO, R. A.; GOODMAN, D. F. A Meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology*, v. 80, n. 5, p.565-579, 1995.
- DE BEM, A. B. *Confiabilidade e validade estatísticas da avaliação docente pelo discente: proposta metodológica e estudo de caso*. 296f. 2004. Tese (Doutorado) – Universidade Federal de Santa Catarina, Florianópolis.
- HORN, B. K. P. *Robot vision*. Cambridge: MIT, 1986.
- HORTALE, V. A. Modelo de avaliação Capes: desejável e necessário, porém, incompleto. *Cadernos de Saúde Pública*, v. 19, n.6, p. 1837-1840, 2003.
- LAKATOS, E. M.; MARCONI, M. A. *Fundamentos da metodologia científica*. 6.ed. São Paulo: Atlas, 2005.
- MATOS, I. H. *Análise de práticas de seleção para o ingresso em cursos de pós-graduação stricto sensu em áreas da Psicologia e Educação*. 1995. Dissertação (Mestrado) – Pontifícia Universidade Católica de São Paulo, São Paulo.
- MEHRENS, W. A.; LEHMANN, I. J. *Measurement and evaluation in education and psychology*. New York: Holt, Rinehart and Winston, 1978.
- ORDEIX, M. D. C. *Seleção para o mestrado na Faculdade de Educação da Universidade de Brasília, turma 1981: uma análise dos critérios e da validade preditiva dos procedimentos adotados*. 1982. Dissertação (Mestrado) – Universidade de Brasília, Brasília.
- PERROCA, M. G.; GAIDZINSKI, R. R. Instrumento de classificação de pacientes de Perroca: teste de confiabilidade pela concordância entre avaliadores – correlação. *Revista da Escola de Enfermagem*. São Paulo, v.36, n.3, p.245-252, 2002.
- POLIT, D. F.; HUNGLER, B. P. *Fundamentos de pesquisa em enfermagem*. 3.ed. Porto Alegre: Artes Médicas, 1995.
- QUELLMALZ, E. S. *The International encyclopedia of educational evaluation*. Oxford: Pergamon, 1990. p. 513-514. Edited by Herbert J. Walberg and Geneva D. Haertel.
- RODRIGUES JR., J. F.; CARNIELLI, B. L.; MOURA, E. A. O Processo de seleção ao Mestrado em Educação: percepções de candidatos. In: ENCONTRO DE PESQUISA EM EDUCAÇÃO DA REGIÃO CENTRO OESTE, 8., 2006, Cuiabá. *Anais...* Cuiabá, 2006. p. 1-19.
- SAORIN, R. N. S. *O Sistema de avaliação da Capes na visão dos gestores dos programas de pós-graduação*. 2009. Dissertação (Mestrado) – Universidade Federal da Paraíba, João Pessoa.
- SPAGNOLO, F.; SOUZA, V. C. O Que mudar na avaliação da Capes? *Revista Brasileira de Pós-Graduação*, v.1, n.2, p. 8-34, nov. 2004.
- STATISTICAL Package for the Social Sciences (SPSS), Versão 11.0. Chicago: SPSS, 2001.
- STUFFLEBEAM, D. L. (Coord.) *Educational evaluation & decision making*. Bloomington: Phi Delta Kappa, 1971.
- UNIVERSIDADE CATÓLICA DE BRASÍLIA. *Edital de abertura das inscrições para a seleção de candidatos ao mestrado em educação da Universidade Católica de Brasília*. Disponível em: <<http://www.ucb/pós-graduação/mestrados>>. Acesso entre: set. e dez. 2006.

Recebido em: fevereiro 2010

Aprovado para publicação em: abril 2010