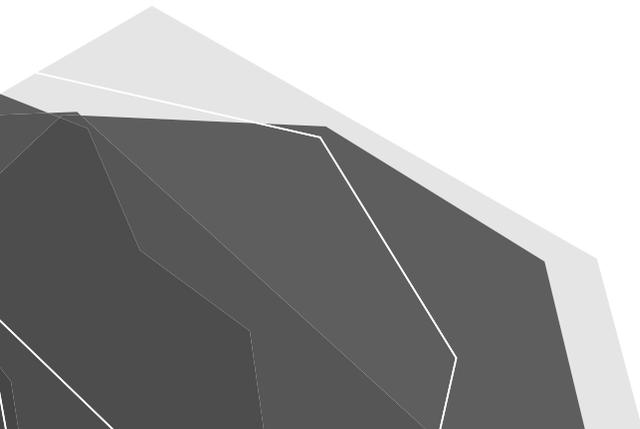


**RE  
SE  
NHA**

review



# PADRÕES PARA EDUCAÇÃO E PSICOLOGIA

STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING

---

**Cristiane Faiad**

**Carlos Manoel Lopes Rodrigues**

*AERA, APA, & NMCE. Standards for Educational And Psychological Testing.*

---

O *Standards*, traduzido como Padrões para Educação e Psicologia, foi publicado pela *American Educational Research Association* (AERA), *American Psychological Association* (APA) e pela *National Council on Measurement in Education* (NCME) como fonte de orientação para construção e uso de testes, na tentativa de promover melhores práticas no uso da medida, subsidiar a qualificação dos mesmos e orientar a busca por critérios psicométricos para adequada interpretação dos escores fornecidos pelos instrumentos. Documentos desta natureza são propostos desde 1954, por diferentes comitês de gerenciamento de normas, embora a versão no padrão do *Standards* tenha sido desenhada em 1999. Já, em 2014, houve a necessidade de que o documento fosse atualizado, sendo esta a última versão disponível. Uma importante consideração é feita no *Standards*, ao explicar que, apesar de diferenças no uso do termo escalas, inventários e testes, este último é utilizado para se referir a todo tipo de instrumento avaliativo. O teste é definido como um procedimento que permite a avaliação de uma amostra de comportamento de um examinando, em determinado domínio, por meio de um processo padronizado. Espera-se, contudo, que o documento seja aplicado para além das medidas utilizadas nos campos da educação e da psicologia, como habilidade, aptidão, realização, atitudes, interesses, personalidade, funcionamento cognitivo e saúde mental, estendendo-se para outras medidas no campo da avaliação. Enquanto documento que estabelece diretrizes, chama a atenção de que contempla orientações importantes e mínimas na construção de medidas, embora não isente olhares mais técnicos, quando se trata de contextos judiciais e de normativas existentes sobre o assunto.

A primeira parte da obra se concentra nas questões relativas aos fundamentos do desenvolvimento e uso dos testes, englobando temas como validade, confiabilidade e equidade. Quanto à validade, o *Standards* apresenta sua importância e uma definição de que “refere-se ao grau em que a evidência e a teoria suportam a interpretação dos escores do teste para o uso proposto pela medida” (p. 9), ou seja, considerando um processo de acúmulo de evidências relevantes que proporcionem uma base científica sólida para a interpretação dos resultados dos testes. Esse é um processo contínuo de responsabilidade de desenvolvedores e usuários dos testes, que se inicia com o planejamento do instrumento e continua por meio de seu desenvolvimento e implementação, envolvendo cuidados com os possíveis vieses que possam emergir por representações inadequadas do construto, ou por aspectos do instrumento, tais como formato do teste, condições de administração, e nível da linguagem. As evidências de validade são organizadas no *Standards* como: a) orientadas ao conteúdo; b) relativas ao processo cognitivo subjacente e ao processo de resposta; c) relativas à estrutura interna e às relações com construtos conceitualmente; d) relacionadas aos critérios; e e) como evidências baseadas nas consequências do teste. O destaque dado para as consequências da testagem alinha-se à preocupação presente em toda a obra quanto à responsabilidade e ao impacto social que o processo de avaliação acarreta.

Quanto à confiabilidade e precisão, parte-se da noção geral de consistência dos resultados através das várias instâncias dos procedimentos de teste, sendo a necessidade de precisão proporcional à importância da interpretação e das decisões decorrentes dos resultados dos testes. São apresentadas formas de análise de confiabilidade e precisão (como, por exemplo, o erro-padrão de medida, a Teoria da Generalizabilidade e a Teoria da Resposta ao Item), mas na prática a avaliação da confiabilidade/precisão é feita utilizando-se diferentes coeficientes, a depender do foco da análise e do modelo de medida em uso. O destaque é dado, na primeira parte, à equidade nos testes, com a proposição de dois novos conceitos: o de acessibilidade e o de *design* universal. Por acessibilidade, entende-se que todos os examinados devem ter oportunidade de realizar o teste, levando em consideração sua posição/condição específica; e o *design* universal compreendido como uma abordagem que busca maximizar a acessibilidade para todos os candidatos que se pretende medir. Especificamente o *design* universal requer construtos precisamente definidos, procedimentos e instruções simples, claros e intuitivos e máxima legibilidade e compreensibilidade. O atendimento deste requisito inclui uma forma

mais globalizante de pensar a construção de instrumentos, considerando questões técnicas e sociais mais amplas.

A segunda parte do *Standards* denominada “operações” compreende recomendações e procedimentos desde o projeto e desenvolvimento do teste, até o desenvolvimento de normas e padrões de interpretação. Esta parte inclui, também, duas sessões específicas quanto aos direitos e deveres dos desenvolvedores e usuários dos testes. Além dos procedimentos e técnicas já reconhecidos, alguns pontos são destacados, como a necessidade de envolver psicometristas na fase de desenvolvimento, avaliar a relação custo-benefício e a utilidade das ações de desenvolvimento dos testes. O próximo destaque é dado à criação de normas e padrões de interpretação dos testes, com a discussão centrada nas formas de redução de vieses de interpretação em prol de decisões com maior generalidade e equidade. Esta preocupação é reforçada no detalhamento dado à definição das formas de aplicação dos testes e no fornecimento de informações detalhadas sobre o teste para o futuro uso adequado. O apontamento dos direitos e deveres dos desenvolvedores e usuários dos testes ressalta essa preocupação ética que perpassa todo o *Standards*.

Por fim, a terceira parte é direcionada à aplicação dos testes em diferentes contextos, sendo destacado o uso no psicodiagnóstico (capítulo 10), no contexto organizacional e de acreditação profissional (capítulo 11), nas avaliações educacionais (capítulo 12) e na avaliação de programas e políticas públicas.

Quando comparada a versão atual com a anterior, de 1999, observa-se que os fundamentos e operações foram destacados para as duas primeiras partes da obra, dedicando assim maior espaço para estas questões. O uso de capítulos sequenciais específicos sobre responsabilidade e direitos de desenvolvedores e usuários de testes também representa uma mudança introduzida nesta versão, já que na versão de 1999 estes tópicos vinham em partes separadas da obra. A discussão sobre equidade, que, na versão antecessora, ocupava a segunda parte da obra, agora permeia todos os capítulos como um fio condutor que alinha toda a preocupação ética e social das organizações autoras. Assim, a versão atual do *Standards* apresenta, além do conjunto de boas práticas na construção e uso de instrumentos de medida, um retrato do desenvolvimento da área ao incorporar inovações técnicas e questões éticas relevantes para o campo. A obra, portanto, não é só uma ótima referência para o processo de construção de medidas, mas também para o usuário final dos instrumentos.

**Cristiane Faiad**

Doutora em Psicologia Social, do Trabalho e das Organizações (PSTO) pela Universidade de Brasília.

Professora da Universidade de Brasília, Brasília (DF), Brasil;  
crisfaiad@gmail.com

**Carlos Manoel Lopes Rodrigues**

Mestre em Psicologia Social, do Trabalho e das Organizações pela Universidade de Brasília.

Professor Assistente do Centro Universitário de Brasília – UNICEUB, Brasília (DF), Brasil;  
prof.carlos.manoel@gmail.com

Resenha submetida em 30/05/2018

Aprovada em 11/06/2018