

## Avaliação Psicológica no Brasil e no Mundo

José Maurício Haas Bueno<sup>1</sup>

<sup>1</sup>Universidade Federal de Pernambuco, PE, Brasil

Evandro Moraes Peixoto<sup>2</sup>

<sup>2</sup>Universidade de Pernambuco, PE, Brasil

**Resumo:** A área de avaliação psicológica tem uma relevância histórica no desenvolvimento da Psicologia como ciência e como profissão, tanto no contexto internacional quanto no nacional. Apesar disso, ao longo de sua história houve momentos tanto de forte crédito quanto de descrédito científico e popular. A compreensão dos fatores que levaram a esses movimentos de altos e baixos é fundamental para o encontro de um ponto de equilíbrio, que permita o desenvolvimento sustentável e contínuo dessa área do conhecimento. Assim, este trabalho teve o objetivo de traçar um paralelo entre a história da avaliação psicológica no Brasil e no mundo, de modo a contribuir para o entendimento dos caminhos percorridos, seus erros e acertos, e possíveis direções futuras. Foram abordados os aspectos éticos e técnicos, bem como ações político-administrativas identificadas como importantes nessa trajetória e seus impactos no desenvolvimento dessa área da Psicologia. Finalmente, são apontadas possíveis direções futuras especialmente em relação ao Sistema de Avaliação de Testes Psicológicos (Satepsi) e à formação em avaliação psicológica. Conclui-se que a área de avaliação psicológica tem sido capaz de superar as críticas recebidas, usando-as como fonte de inspiração para a realização de pesquisas, desenvolvimento de novas tecnologias e práticas mais afinadas com os direitos humanos, de modo a contribuir com o fortalecimento científico e profissional da própria Psicologia e de seu compromisso com a sociedade.

**Palavras-chave:** Avaliação Psicológica, História da Psicologia, Psicometria, Ética, Direitos Humanos.

---

## Psychological Assessment in Brazil and in the World

**Abstract:** Psychological assessment has had a historical relevance in the development of psychology as science and as a profession, both in international and in national contexts. Nevertheless, there have been times of both strong scientific and popular credit and discredit throughout its history. Understanding the factors that led to these ups and downs is paramount to finding the point of balance toward a sustainable and continuous development of this area of knowledge. Thus, the aim of this study was to draw a parallel between the history of psychological assessment in Brazil and in the world, in order to contribute to the understanding of the paths covered, the mistakes and successes, and the possible future directions. The ethical and technical aspects, as well as administrative and political actions identified as important in this trajectory and their impacts on the development of this area of psychology were discussed. Then, possible future directions are pointed out, especially in relation to the Psychological Tests Evaluation System (SATEPSI) and to the professional qualification in psychological assessment. We concluded that the researchers on psychological assessment have been able to take advantage of the criticisms as a source of inspiration for researches planning, development of new technologies and professional practices more attuned to the human rights, so as to contributing to scientific and professional strengthening of psychology and its commitment to society.

**Keywords:** Psychological Assessment, Psychology History, Psychometric, Ethics, Human Rights.

## Evaluación Psicológica en Brasil y en el Mundo

**Resumen:** El área de evaluación psicológica tiene una relevancia histórica en el desarrollo de la psicología como ciencia y como profesión, tanto en el contexto internacional como en el nacional. A pesar de ello, a lo largo de su historia hubo momentos tanto de fuerte crédito y de descrédito científico y popular. La comprensión de los factores que llevaron a estos movimientos de altibajos es fundamental para el encuentro de un punto de equilibrio que permita el desarrollo sostenible y continuo de esa área del conocimiento. Así, este trabajo tuvo el objetivo de trazar un paralelo entre la historia de la evaluación psicológica en Brasil y en el mundo, de modo a contribuir para el entendimiento de los caminos recorridos, sus errores y aciertos, y posibles direcciones futuras. Se abordaron los aspectos éticos y técnicos, así como acciones político-administrativas identificadas como importantes en esa trayectoria y sus impactos en el desarrollo de esa área de la psicología. Finalmente, se señalan posibles direcciones futuras especialmente en relación al Sistema de Evaluación de Pruebas Psicológicas (SATEPSI) y a la formación en evaluación psicológica. Se concluye que el área de evaluación psicológica ha sido capaz de superar las críticas recibidas, usándolas como fuente de inspiración para la realización de investigaciones, desarrollo de nuevas tecnologías y prácticas más afinadas con los derechos humanos, para contribuir con el fortalecimiento científico y profesional de la propia psicología y de su compromiso con la sociedad.

**Palabras clave:** Evaluación Psicológica, Historia de la Psicología, Psicometría, Ética, Derechos Humanos.

### Introdução

A área de avaliação psicológica tem uma relevância histórica no desenvolvimento da Psicologia como ciência e como profissão, tanto no contexto internacional quanto no nacional. No Brasil, essa área foi incluída na própria Lei Federal nº 4.119 (1962), que regulamentou a profissão de psicólogo no país e, entre outras coisas, estabelecia apenas uma função como privativa do psicólogo: a utilização de métodos e técnicas psicológicas para fins de diagnóstico psicológico, orientação e seleção profissional, orientação psicopedagógica e solução de problemas de ajustamento. Com base nisso, o Conselho Federal de Psicologia (CFP) considera que os testes psicológicos se enquadram nessa definição e, por isso, se constituem em métodos ou técnicas de uso privativo dos psicólogos (Resolução Nº 009, de 25 de abril de 2018).

Embora se pudesse supor que isso levaria a Psicologia brasileira a conferir grande importância à formação, pesquisa e atuação nessa área, o que ainda hoje se observa é um movimento contraditório, em que a testagem psicológica é apresentada aos estudantes como a única prática privativa dos psicólogos e como produtora de exclusão social (Patto, 1997).

Essa situação sofreu uma alteração significativa por volta da virada do século, em decorrência de uma série de ações políticas no âmbito da avaliação psicológica no Brasil, que impactaram positivamente o desenvolvimento dessa área até os dias atuais. Ainda assim, muito se tem para fazer e a continuidade desse desenvolvimento passa pela necessidade reafirmar o compromisso com a sociedade, estimulando, por um lado, o desenvolvimento de práticas adequadas ao contexto brasileiro e, por outro, a sintonia com as tendências tecnológicas e sociais observadas em outras partes do mundo. Assim, este trabalho teve o objetivo de traçar um paralelo entre a história da avaliação psicológica no Brasil e no mundo, de modo a contribuir para o entendimento dos caminhos percorridos, seus tropeços e acertos, e possíveis direções futuras.

### Aspectos históricos: da fundação ao abandono... e a reconstrução!

As principais dificuldades enfrentadas na área de avaliação psicológica durante as quatro últimas décadas do século XX no Brasil, foram, em grande parte, decorrentes de dois fatores: 1) confusão entre os conceitos de avaliação psicológica e testagem psicológica

e 2) má qualidade psicométrica dos instrumentos. Esses fatores não surgiram no contexto brasileiro, ao contrário, são inerentes ao desenvolvimento das técnicas de avaliação psicológica, especialmente dos testes, no cenário internacional.

A confusão entre avaliação e testagem psicológica foi parte do próprio desenvolvimento da Psicologia como ciência. Cattell (1890), por exemplo, apontava o desenvolvimento de instrumentos de medida e a criação de laboratórios experimentais como caminhos pelos quais a Psicologia poderia conquistar o *status* científico. De fato, o início do século XX viu o desenvolvimento de testes psicológicos resultar em grandes avanços teóricos (Spearman, 1904) e práticos Binet (1904) no campo da inteligência, por exemplo, com a proposição do modelo do fator geral e de um teste capaz de identificar crianças com dificuldades de aprendizagem, respectivamente. Essas ocorrências contribuíram significativamente para o estabelecimento da relevância científica da Psicologia como área do conhecimento e do seu compromisso com a sociedade. Fortaleceu-se assim a ideia de que os testes permitiam avaliar aspectos psicológicos de forma comparável ao modelo médico de atendimento, no qual os procedimentos padronizados de diagnóstico (exames), isentos de subjetividade, se constituíam não em *um*, mas *no* método preferencial de diagnóstico (Ocampo, Arzeno & Piccolo, 1981).

Posteriormente, com o advento da II Grande Guerra, estabeleceu-se um conflito entre a necessidade de testes para a seleção de soldados para a guerra e o estágio ainda inicial das bases psicométricas da testagem naquele momento. O resultado desse conflito foi a utilização de instrumentos de avaliação com qualidade psicométrica duvidosa. Assim, os testes, que inicialmente haviam contribuído para o estabelecimento da Psicologia como ciência, em decorrência de seu uso como única fonte de informação para a realização de diagnósticos e de suas deficiências psicométricas, que ainda não eram completamente conhecidas (muito menos controladas) pelos cientistas, passaram a ser questionados quanto a sua eficácia. E esse questionamento recaiu sobre a própria capacidade da Psicologia em dar as respostas esperadas pela sociedade (Bueno & Ricarte, 2017).

É nesse cenário mundial que ocorre a implantação do Instituto de Seleção e Orientação Profissional (ISOP), em 1947, no Rio de Janeiro, cujo objetivo era realizar processos de seleção de pessoas baseados em

evidências científicas, compreendidas à época como tomadas de decisão baseadas nos resultados de testes psicológicos (Noronha, & Reppold, 2010). Nesse sentido, a avaliação psicológica no Brasil ainda refletia a confiança nos resultados dos testes observada nos processos americanos de seleção de pessoas para os esforços de guerra, mas contribuiu significativamente para a regulamentação da profissão de psicólogo no Brasil, ocorrido em 21 de agosto de 1962, pela Lei nº 4.119.

No entanto, esse cenário favorável logo se deterioraria em razão de dois fatores: 1) a criação acelerada de cursos de Psicologia e a falta de professores qualificados para o ensino, que culminaram na deficiência de formação, especialmente na área de avaliação psicológica, 2) o surgimento de novas correntes de pensamento, como a humanista e a sócio-histórica, que eram reativas à quantificação e ao positivismo e suas manifestações (Pasquali, & Alchieri, 2011). As críticas eram de que os testes serviam mais para rotular e estigmatizar as pessoas do que para ajudá-las a se desenvolver, pelo que propunham o abandono dessa prática que consideravam nociva à própria imagem da Psicologia (Noronha, 2002; Patto, 1997).

Nesse contexto, a área de avaliação psicológica como um todo foi se deteriorando, tanto na formação de profissionais quanto no desenvolvimento de instrumentos mais apropriados. A formação em avaliação psicológica, por exemplo, foi se reduzindo ao questionamento da eficácia dos testes, à afirmação de seu caráter excludente ou ao ensino dos procedimentos de aplicação, contagem de pontos e interpretação estatística (e não psicológica) do resultado. Um dos reflexos desse tipo de formação, foi a diminuição do desenvolvimento psicométrico dos testes (adaptação cultural, estudos de validade etc.) (Gouveia, 2009), o que resultou na comercialização de um conjunto de testes com propriedades psicométricas desconhecidas ou não informadas no manual do instrumento e em um elevado número de processos éticos no CFP e na justiça, contrários aos resultados decorrentes de avaliação psicológica.

No contexto internacional, as críticas foram tomadas como um impulso para a realização de pesquisas e para o desenvolvimento das práticas de avaliação e de testagem. O aprimoramento de técnicas estatísticas, como a análise fatorial, por exemplo, foram fundamentais para o desenvolvimento de teorias importantes da inteligência (Cattell, 1943;

Thurstone, 1931; 1934; 1937; 1938) e da personalidade (Cattell, 1946; 1947), estabelecendo as bases para o desenvolvimento de instrumentos fundamentos no conceito de traço latente. Além disso, a criação da *Psychometric Society*, em 1935, e de seu periódico *Psychometrika*, em 1936, foram precursores de muitas outras sociedades e revistas científicas dedicadas ao tema, que continuam ativas até os dias atuais (Bueno, & Ricarte, 2017).

Cenário semelhante só foi surgir com maior força no Brasil, nos anos 1990, quando se observou uma movimentação oriunda de professores e pesquisadores descontentes com a situação da avaliação psicológica, tanto no âmbito da formação quanto da pesquisa. Por isso, organizaram uma série de eventos, nos quais invariavelmente se discutia, além dos resultados de pesquisas, o papel das disciplinas de avaliação psicológica na formação, políticas para o desenvolvimento da área no Brasil, entre outros temas. São exemplos de eventos realizados naquele período o I e o II Encontro de Técnicas do Exame Psicológico: Ensino, Pesquisa e Aplicações, na Universidade de São Paulo, o I Encontro Nacional sobre Testes Psicológicos (Porto Alegre, UFRGS / PUC-RS), o Encontro Mineiro de Avaliação Psicológica (Belo Horizonte, UFMG / PUC-MG / Centro Universitário Newton Paiva / Fumec) e o Encontro da Sociedade Brasileira de Rorschach e outros Métodos Projetivos (Ribeirão Preto, USP-RP) (Primi, 2010). Dois importantes resultados desses encontros foram a criação da Sociedade Brasileira de Rorschach e Outros Métodos Projetivos (SBRo)<sup>1</sup> (<http://www.asbro.org.br/>), em 1993, e do Instituto Brasileiro de Avaliação Psicológica (IBAP) (<http://www.ibapnet.org.br/>), em 1997.

Essas associações profissionais permanecem ativas atualmente e organizam congressos bianuais, intercalados, de modo que todo ano há um congresso de abrangência nacional na área de avaliação psicológica. Além disso, o IBAP edita a revista *Avaliação Psicológica*, ([http://pepsic.bvsalud.org/scielo.php?script=sci\\_serial&pid=1677-0471&lng=pt&nrm=iso](http://pepsic.bvsalud.org/scielo.php?script=sci_serial&pid=1677-0471&lng=pt&nrm=iso)), atualmente com periodicidade quadrimestral e avaliada pelo Qualis da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) como A2 (<http://www.capes.gov.br/acesoainformacao/perguntas-frequentes/avaliacao-da-pos-graduacao/7422-qualis>).

Em meados dos anos 1990 havia quatro laboratórios pioneiros de pesquisa em avaliação psicológica,

situados em importantes universidades brasileiras: o Laboratório de Pesquisa em Avaliação e Medida (LAB-PAM), na Universidade de Brasília, fundado em 1988 pelo Prof. Dr. Luiz Pasquali; o Centro de Pesquisas em Psicodiagnóstico (CPP), no campus de Ribeirão Preto da Universidade de São Paulo, fundado em 1975 pelo Prof. Dr. André Jacquemin; o Laboratório de Avaliação e Medidas Psicológicas (LAMP), na Pontifícia Universidade Católica, na cidade de Campinas, fundado em 1994 e coordenado pela Prof.<sup>a</sup> Dr.<sup>a</sup> Solange Wechsler e o Laboratório de Mensuração (LM), na Universidade Federal do Rio Grande do Sul, fundado em 1988 pelo Prof. Dr. Cláudio Hutz. Esses laboratórios e professores deram grande impulso à pesquisa e à formação especializada em avaliação psicológica no Brasil, fundaram linhas de pesquisa e assumiram papéis de destaque na organização dos eventos e associações citadas anteriormente (Bueno, & Ricarte, 2017).

Em 1996, o relatório do Comitê Assessor de Psicologia do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), apontou para a área de Fundamentos e Medidas Psicológicas como uma das cinco que mereciam mais investimentos à época. Essa indicação foi levada em consideração para a criação de um Programa de Pós-graduação com área de concentração em avaliação psicológica, na Universidade São Francisco, no ano 2000, que é avaliado atualmente com nível de excelência internacional pela Capes. A implantação desse programa de pós-graduação deu grande impulso aos esforços dos centros pioneiros descritos anteriormente quanto à formação, produção de conhecimentos e uso de modernas técnicas de avaliação psicológica para resolução de problemas sociais.

Esse conjunto de ocorrências teve impacto nas políticas implementadas pelo CFP, que observava, por um lado, o esforço desses profissionais no sentido da construção de boas práticas de avaliação psicológica, e, por outro, o grande número de processos éticos em decorrência das más práticas em avaliação psicológica, descritas anteriormente. Por isso, no início dos anos 2000, o CFP instituiu a Comissão Consultiva em Avaliação Psicológica, formada por pesquisadores e representantes ligados à ASBRo e ao IBAP, com o objetivo de propor políticas que ajudassem a melhorar a qualidade das práticas em avaliação psicológica (*Resolução Nº 25, 2001*).

<sup>1</sup> Posteriormente, a SBRo viria a se chamar Associação Brasileira de Rorschach e Outros Métodos Projetivos (ASBRo), denominação que permanece até o presente momento.

A proposta apresentada pelo CFP baseou-se no fortalecimento do compromisso da Psicologia com a sociedade, por meio de um sistema de monitoramento contínuo da qualidade técnica dos instrumentos de avaliação utilizados pelos psicólogos. Assim, foi implantado o Sistema de Avaliação de Testes Psicológicos (Satepsi) (*Resolução Nº 25, 2001*), que, segundo Mansur-Alves, Silva e Fernandes (2016), se constituiu em um dos fatores mais importantes para a retomada do desenvolvimento da área de avaliação psicológica no Brasil. De fato, esse sistema teve grande importância tanto prática, porque estabelecia critérios para o uso de testes, quanto simbólica, pois a principal instituição à qual os psicólogos estão filiados implementava uma política nacional em relação ao uso de testes. A próxima seção apresenta mais detalhadamente esse sistema.

### **Sistema de Avaliação de Testes Psicológicos**

O Satepsi foi proposto pelo CFP em 2001, entre outros motivos, pela necessidade de aprimorar os instrumentos e procedimentos técnicos dos psicólogos e de garantir uma prestação de serviços com qualidade técnica e ética aos seus usuários (*Resolução Nº 25, 2001*). Posteriormente, o sistema foi atualizado pela Resolução nº 02/2003, e, mais recentemente, pela Resolução nº 09/2018, que continua estabelecendo as diretrizes básicas para a realização de avaliação psicológica no exercício profissional da psicóloga e do psicólogo, incluindo os procedimentos de submissão e de avaliação dos testes e/ou suas atualizações no sistema, e acrescenta questões relacionadas com a justiça e proteção dos direitos humanos na avaliação psicológica (*Resolução Nº 09, 2018*).

A avaliação dos instrumentos que são submetidos ao sistema é facilitada pelo Formulário de Avaliação da Qualidade de Testes Psicológicos, constante como Anexo 1 da resolução, que sistematiza a coleta de informações sobre seis aspectos que devem constar nos manuais dos instrumentos: 1) qualidade do manual; 2) fundamentação teórica, 3) análise de itens; 4) precisão; 5) validade; e 6) sistema de correção e interpretação dos resultados. Esses itens são avaliados em todos os testes considerados psicológicos, exceto para os projetivos, para os quais a análise de itens não se aplica. Ao final da avaliação de cada tópico, o avaliador emite um parecer geral favorável ou desfavorável. Para que o parecer seja favorável o instrumento

deve receber conceitos equivalentes a excelente ou bom em todas as áreas avaliadas. Antes de entrarem em vigor, os pareceres passam pela avaliação da Comissão Consultiva em Avaliação Psicológica e do Plenário do CFP (*Resolução Nº 09, 2018*).

Uma iniciativa internacional semelhante ao Satepsi é a do *COnsensus-based Standards for the selection of health Measurement INstruments* (Cosmin). O Cosmin, mais do que uma medida reguladora, foi proposto como um marco referencial, no qual os pesquisadores pudessem se basear para construir e/ou avaliar a qualidade de instrumentos de medida na área da saúde. Seu uso é recomendado para revisões sistemáticas de instrumentos de medida, seleção de instrumentos de medida, identificação da necessidade de mais estudos sobre as propriedades de medida dos instrumentos, elaboração de estudos sobre propriedades de medida, avaliar a qualidade de manuscritos sobre instrumentos de medida submetidos para publicação, entre outros (Mokkink et al., 2012).

Seu surgimento pode ser localizado em 2005, num esforço pela reunião de especialistas com o objetivo de estabelecer princípios, procedimentos e critérios comuns, que servissem como base (taxonômica, terminológica e paramétrica) para o desenvolvimento de instrumentos e de boas práticas de medida em ciências da saúde (Mokkink et al., 2006). Para isso foi realizado um estudo com a metodologia Delphi (Jorm, 2015), que, basicamente, procura encontrar consenso entre especialistas em uma determinada área. Este estudo resultou na identificação consensual das propriedades de medida a serem consideradas na avaliação de um instrumento (Mokkink et al., 2010).

Foram identificados três domínios de informações relevantes sobre os instrumentos de medida: fidedignidade, validade e a responsividade, sendo que dentro de cada domínio são consideradas vários tipos de propriedades da medida. O domínio da fidedignidade se refere ao grau com que um instrumento está livre de erros de medida. Isto pode ser avaliado pela intercorrelação entre os itens (consistência interna), pela proporção da variância total que representa diferenças verdadeiras entre os pacientes, ou pelo erro sistemático e aleatório da pontuação de um paciente que não é atribuído a mudanças verdadeiras no construto sob consideração (erros de medida) (Mokkink et al., 2012).

O domínio da validade é compreendido como o grau com que um instrumento mede aquilo que ele se propõe a medir. Uma das formas de se investigar

esse domínio é por meio do grau com que um instrumento reflete (validade de conteúdo) ou parece refletir (validade de face) o construto a ser mensurado. Outra forma se refere ao grau com que o instrumento é consistente com as hipóteses que se pode derivar do construto (validade de construto), que pode ser investigado por meio da dimensionalidade do instrumento (validade estrutural), do teste de hipóteses (validade de construto propriamente dita) e da semelhança de funcionamento de itens traduzidos de uma língua/cultura para outra (validade transcultural). Outra forma, ainda, de investigação da validade é verificar o grau com que os escores de um instrumento refletem adequadamente um “padrão ouro” (validade de critério) (Mokkink et al., 2012).

A responsividade, por sua vez, é definida como a capacidade de um instrumento em detectar mudanças ao longo do tempo em um determinado construto. Ao lado desses três domínios mais métricos, há ainda o domínio da interpretabilidade, que não é exatamente uma propriedade da medida, mas é importante para a compreensão da medida, uma vez que se refere ao grau com que se consegue atribuir um significado qualitativo aos resultados quantitativos obtidos (escores) (Mokkink et al., 2012).

De forma semelhante ao Satepsi, os itens a serem avaliados foram operacionalizados numa lista de verificação (*checklist*), organizada em 12 seções. As primeiras nove seções, identificados por letras, se referem às propriedades de medida: A – Consistência interna, B – Fidedignidade, C – Erro de mensuração, D – Validade de Conteúdo, E – Validade estrutural, F – Testagem de Hipóteses, G – Validade Transcultural, H – Validade de critério e I – Responsividade. A seção J se refere aos estudos de interpretabilidade. A seção seguinte (TRI) se refere a estudos realizados com auxílio da Teoria de Resposta ao Item (TRI). Por fim, a última seção (Generalização) apresenta os requisitos para avaliação da possibilidade de generalização dos resultados para outras amostras e população de uma forma geral (Mokkink et al., 2010; 2012).

Essa lista de verificação contém um total de 129 itens, mas foi elaborada para ser usada de forma modular. Isto é, pode-se utilizar qualquer de suas seções de forma independente, de acordo com o propósito de sua avaliação. Assim, os autores descrevem um sistema de quatro passos para sua utilização: 1) Determinação de quais seções precisam ser respondidas (entre as nove seções iniciais, identifica-

das pelas letras de A a I); 2) Preenchimento da seção sobre utilização da TRI (caso tenha sido utilizada), 3) Preenchimento das seções assinaladas no passo 1; e 4) Preenchimento das informações sobre generalização dos resultados, para cada seção assinalada no primeiro passo. Algumas das seções da lista de verificação podem ter que ser preenchidas mais de uma vez, conforme os dados avaliados se apresentem. Por exemplo, dados de fidedignidade apresentados para diferentes amostras da população, requerem o preenchimento da seção de B (Fidedignidade) tantas vezes quantos forem as amostras consideradas no estudo (Mokkink et al., 2012).

Há duas formas de avaliação dos itens: 1) com respostas dicotômicas, que indicam a presença (SIM) ou ausência (NÃO) de determinada informação sobre o instrumento (Mokkink et al., 2012), e 2) com quatro categorias de classificação: excelente, bom, razoável e ruim (Terwee et al., 2012). O segundo método é empregado para cada seção do COSMIN, sendo que o critério adotado é para avaliação é o da pior pontuação. Por exemplo, se um item da E (validade estrutural) é avaliado como “ruim”, então toda a seção é avaliada como “ruim”. Nesse sistema, a avaliação “ruim” em qualquer dos itens é considerada uma “falha fatal” (Terwee et al., 2012).

Não foram encontrados estudos de impacto do Cosmin nas práticas profissionais e/ou de investigação, mas uma busca realizada em junho/2018 no Portal Periódicos da Capes (<http://www.periodicos.capes.gov.br/>) com o termo “Cosmin” no título ou no assunto, retornou 270 referências publicadas desde 2005. A maior parte desses estudos foi publicada nos últimos cinco anos, o que parece indicar sua crescente utilização como referência para a construção e avaliação da qualidade de instrumentos de medida na área da saúde.

Por ser um instrumento mais regulador sobre a prática profissional do psicólogo, O Satepsi teve um impacto bastante importante para o desenvolvimento da área de avaliação psicológica no Brasil. Por exemplo, em 2002, logo após sua implantação, dos 91 testes submetidos ao sistema, mais da metade (61,5%) foi considerada “desfavorável” para uso profissional (Noronha, Primi, & Alchieri, 2004). Houve mobilizações dos próprios profissionais contrários à implantação do Satepsi e até reportagens em importantes meios de comunicação impressos (jornais e revistas) da época, questionando a eficácia dos instrumentos de avaliação psicológica. O CFP sofreu foi acionado

na justiça pelas editoras, mas obteve sentença favorável em todas as decisões. Com o tempo, essas reações negativas diminuíram, os testes foram melhorando de qualidade e a percepção tornou-se mais positiva, embora nunca isenta de críticas, muitas delas construtivas e que deveriam ser consideradas para evolução do sistema e da qualidade da avaliação psicológica no Brasil.

Uma novidade proporcionada pela Resolução do CFP nº 09/2018 foi incorporar na própria resolução que trata do Satepsi, uma seção relacionada com a justiça e proteção dos direitos humanos na avaliação psicológica. Nessa seção, o CFP alerta os psicólogos e psicólogas contra certas práticas na avaliação psicológica, que possam caracterizar negligência, preconceito, exploração, violência, crueldade ou opressão; induzir a convicções políticas, filosóficas, morais, ideológicas, religiosas, raciais, de orientação sexual e identidade de gênero; ou favorecer o uso de conhecimento da ciência psicológica e normatizar a utilização de práticas psicológicas como instrumentos de castigo, tortura ou qualquer forma de violência.

De fato, a Avaliação Psicológica e a testagem psicológica podem ter efeitos importantes na vida das pessoas, podendo esses efeitos serem positivos ou negativos, a depender de como estes procedimentos são realizados (Koocher, & Keith-Spiegel, 2016). A próxima seção tratará de questões éticas envolvidas na avaliação psicológica.

### **Questões éticas na avaliação psicológica**

Segundo Hutz (2002) indivíduos e instituições se beneficiam do processo de avaliação psicológica quanto este ajuda os indivíduos a atingirem seus objetivos. Nessa direção a avaliação psicológica pode contribuir para melhora da qualidade de vida das pessoas através da correta identificação de seus potenciais e fragilidades e, portanto, na proposição da intervenção mais adequada a cada situação. Por outro lado, esse mesmo procedimento possui um potencial de causar danos às vidas das pessoas, em especial nos casos em que não se pode contar com métodos, técnicas e instrumentos apropriados a realidade dos sujeitos submetidos a avaliação, bem como em situações de má qualificação do psicólogo que realiza a avaliação psicológica.

Exemplos de malefícios causados pela má utilização da avaliação ou dos testes psicológicos são comuns na literatura e podem ser observados em diferentes

lugares ao redor do mundo. Nesta direção, Laher e Cockcroft (2016) descreveram a maneira como esses procedimentos e técnicas foram utilizados para justificar o regime do *Apartheid* na África do Sul. Segundo os autores, os testes psicológicos foram desenvolvidos na África do Sul em um contexto em que a distribuição de recursos econômicos e educacional e o acesso à liberdade eram muito injustos. Estes instrumentos eram desenvolvidos e normatizados com foco na avaliação das pessoas brancas com acesso à educação, recursos socioeconômicos e livres, e posteriormente administrados em negros que não tinham direitos a tais privilégios. Assim, a constatação de inferioridade das pessoas negras nos processos avaliativos e de testagem eram utilizados para justificar a superioridade intelectual dos brancos, a exploração do trabalho das pessoas negras e, portanto, garantir os privilégios de diferentes ordens para um grupo em detrimento do outro.

Certamente, o mau uso da avaliação psicológica e da testagem não serviu de suporte às injustiças sociais apenas no regime do *Apartheid*, também é possível observar ao longo da história americana o uso inadequado dessas ferramentas. Chama atenção as maneiras controversas como os testes eram administrados com objetivo de comparações entre grupos sociais e, por fim, a reafirmação de um grupo sobre outros, em especial o enaltecimento da classe média branca norte-americana. Sue, Arredondo e McDavis (1992) descrevem os diferentes modelos teóricos alicerçados nas interpretações entre resultados apresentados por brancos americanos em comparação a outras minorias étnicas. O primeiro é o modelo de inferioridade e patologia, que justificava as diferenças nos escores em função de deficiências genéticas e biológicas apresentadas pelas minorias étnicas. O modelo de privação cultural, que defendia que estas diferenças ocorriam em função das crenças, valores e práticas culturais ligadas a essas minorias que, portanto, eram menos valorizadas por brancos americanos. Felizmente, ao longo da história esses modelos foram revistos quanto aos seus conteúdos racistas e insensibilidade cultural, dando lugar a compreensões mais sensatas, como o modelo culturalmente diferente, que não compreendia as diferenças culturais entre brancos outros grupos étnicos como uma indicação da patologia ou inferioridade de qualquer grupo, mais do que isso, defendia a importância dessas diferenças na composição da identidade desses grupos e a riquezas das diversidades culturais.

De acordo com Leong e Park (2016), esse movimento em direção a sensibilidade cultural influenciou diretamente avaliação psicológica e as formas interpretação dos escores dos testes, haja vista a maior atenção direcionada a questões como direitos humanos, equidades entre diferentes grupos étnicos e populações na realização desses processos. A principal expressão desse cuidado é verificada nos *Standards for educational and psychological testing*, proposto pelas principais instituições relacionadas à testagem psicológica na América do Norte: *American Educational Research Association*, *American Psychological Association*, e *National Council on Measurement in Education*, onde são enfatizadas questões como justiça, beneficência e não maleficência ao se propor a comparação entre grupos e populações.

Essa preocupação influenciou também a compreensão do conceito de validade dos testes conforme proposto nos *Standards* (AERA, APA & NCME, 2014). Considerada um ponto fundamental relativo à cientificidade dos testes, a validade refere-se ao grau em que teoria e evidências empíricas sustentam as interpretações dos escores do teste. Quando o teste é utilizado e interpretado de diferentes maneiras ou frente a diferentes grupos ou contextos, exige-se que cada uma das interpretações pretendidas seja antecipada por evidência de validade. Ou seja, não se objetiva a validação do teste em si, mas das possíveis interpretações propostas a partir de seus escores. Segundo AERA et al. (2014), “validação pode ser vista como o desenvolvimento de um argumento cientificamente sólido para apoiar a interpretação proposta a partir dos resultados dos testes, bem como a relevância de sua utilização” (p. 9). Neste caso, o processo de validação de um instrumento envolve o acúmulo de evidências a fim de fornecer bases científicas para a interpretação dos escores obtidos por um sujeito avaliado através daquele teste.

Segundo Primi, Muniz e Nunes (2009), esta percepção contemporânea de validade teve como base as contribuições de Messick (1989), com o questionamento da definição tripartite do conceito de validade: validade de conteúdo, validade de critério e validade de construto, pois segundo Messick mesmo os tipos de validade de conteúdo e critério apresentavam informações referentes ao construto, ou seja, evidências de diferentes naturezas referentes ao sentido de interpretação dos escores do teste e, portanto, elementos sobre o quão bem o teste mensura determinado construto.

Segundo o autor, tais informações são denominadas “fontes de evidência de validade”. De acordo com AERA et al. (2014), são definidas cinco fontes de evidência de validade: a) baseadas no conteúdo, que apresentam informações referentes a representatividade dos itens quanto sua consistência e abrangência na avaliação do construto alvo; b) baseadas no processo de resposta, que fornecem informações referentes aos processos mentais envolvidos na resolução das tarefas apresentadas pelos itens do teste; c) baseadas na estrutura interna, que fornecem informações sobre a estrutura das correlações entre os itens que avaliam o mesmo construto, e estrutura de correlações entre subtestes que avaliam construtos diferentes; d) baseadas na relação com outras variáveis, levanta dados sobre os padrões de correlações entre os escores dos testes e outras variáveis medindo o mesmo construto ou construtos relacionados (convergência), medindo construtos diferentes (divergência). Também apresente informações sobre a capacidade preditiva do teste (critério externo); e) baseadas nas consequências da testagem, que examina as consequências sociais intencionais e não intencionais do uso do teste para verificar se sua utilização está surtindo os efeitos desejados, de acordo com o propósito para qual foi desenvolvido.

Ainda quando último tipo de evidências referido nos *Standards* alguns esclarecimentos se fazem necessários (haja vista o objetivo deste tópico, a discussão de questões éticas ligadas a avaliação psicológica). Embora se discuta de quem seria de fato tais responsabilidades, o construtor e os psicólogos que utilizarão o instrumento não podem isentar-se dessa responsabilidade. Observa-se que tais procedimentos estão alinhados às preocupações com justiça e equidade na comparação entre grupos e populações. Nessa direção, fica claro a responsabilidade do psicólogo que irá escolher um instrumento para compor um processo de avaliação psicológica, uma vez que será necessário acessar as evidências de validade frente o grupo de interesse e ao contexto onde o instrumento será utilizado, para que se possa criar bases de sustentação e, portanto, validar as interpretações psicológicas dos escores observados na testagem.

Essas novas perspectivas exigem que os profissionais da Psicologia estejam atentos a temas como direitos humanos, equidade e justiça, e que sejam formados adequadamente no que diz respeito a avaliação psicológica e testagem psicológica, pois será a partir dos resultados obtidos através desses procedi-



mentos que intervenções psicológicas serão propostas. Contudo, importantes lacunas têm sido apontadas por pesquisadores em relação a formação em Psicologia na realidade brasileira (Hutz, 2002; Zaia, Oliveira & Nakano, 2018).

Já em 2002 Hutz destacava a necessidade de maior qualificação do profissional de Psicologia quanto a avaliação psicológica e psicométria, pois a falta de treinamento adequado pode levar os psicólogos a usarem os instrumentos inadequadamente, sem o entendimento claro do que estão fazendo, causando malefícios aos sujeitos submetidos a avaliação e, portanto, cometendo importantes faltas éticas (Hutz, 2015). Em pesquisa mais recente, Zaia et al. (2018) analisaram os processos éticos publicados no Jornal do Conselho Federal de Psicologia no período de 2004 a 2016, e verificaram que nos três últimos números da revista 60% dos processos estavam relacionados a avaliação psicológica, dentre os quais se destacavam as queixas relacionadas a má elaboração de laudos psicológicos, má utilização de testes psicológicos e irregularidade na avaliação psicológica. Diante das observações as autoras destacam a importância de maiores investimentos na formação do psicólogo em relação a avaliação psicológica e a testagem, haja vista os resultados nefastos da imperícia e, portanto, má conduta profissional nesta área da Psicologia.

### **Qual o impacto das mudanças na área de avaliação psicológica no Brasil?**

A partir dos anos 90 foram observadas uma série de ações que proporcionaram a reflexão sobre práticas inadequadas, a proposição de alternativas e, principalmente, a retomada do desenvolvimento da avaliação psicológica no Brasil. Os principais eventos que impulsionaram essa retomada, em que também se constituem nos primeiros resultados dos esforços para o fortalecimento da área, foram a criação das associações de profissionais ligadas à avaliação psicológica (IBAP e AsBRo), a implantação de um programa de pós-graduação com área de concentração em avaliação psicológica e de diversos laboratórios e linhas de pesquisa em outros programas espalhados pelo país, a edição de uma revista científica especializada em avaliação psicológica e de outras revistas que dedicavam uma seção de suas edições a essa área e a implantação do Sistema de Avaliação de Testes Psicológicos pelo CFP. Embora não seja possível especificar a porcentagem de contribuição de cada uma dessas

ocorrências, há um conjunto de informações disponíveis do qual se pode inferir o impacto desse conjunto de ações na área de avaliação psicológica no Brasil.

Por exemplo, três estudos observaram um aumento significativo do número de resumos apresentados em congressos (Joly, Silva, Nunes, & Souza, 2007), de dissertações e teses defendidas (Anache, & Corrêa, 2010), de artigos científicos publicados (Barroso, 2010; Mansur-Alves et al., 2016) na área de avaliação psicológica a partir do ano 2000. Concomitantemente, o número de testes cadastrados no Satepsi quase que duplicou entre 2004 e 2010 (Primi, 2010). Esses dados sugerem que houve um maior investimento no desenvolvimento e no aperfeiçoamento de instrumentos de avaliação, o que melhorou sua qualidade técnica, além do crescimento do mercado de testes e da disseminação de seus parâmetros psicométricos de qualidade (Primi, 2010).

Além disso, essas publicações se expandiram geograficamente para além dos estados da região sudeste e da clínica para diversas outras áreas de aplicação da avaliação psicológica, como a Psicologia organizacional e do trabalho, a orientação vocacional, a Psicologia do trânsito e a Psicologia forense (Mansur-Alves et al., 2016). Joly et al. (2007) também observaram um aumento do número de universidades representadas nos congressos da área, que passaram a ser majoritariamente públicas em vez de privadas. Bueno, Amorim e Albuquerque (2017) apontam a abertura de concursos nas universidades públicas para a contratação de professores especializados em avaliação psicológica como um dos fatores responsáveis por essa mudança.

Esses indicadores revelam um impacto positivo dos esforços de melhoria na área de avaliação psicológica. No entanto, esse impacto se concentra mais na área da produção de conhecimentos (pesquisa) e na formação de profissionais qualificados em nível de pós-graduação. Mas ainda há muito o que fazer para (a continuidade do) o desenvolvimento da área, tanto no que se refere a questões técnicas e de formação profissional, quanto de questões políticas e de inclusão social e respeito aos direitos humanos. Esses aspectos serão discutidos na próxima seção.

### **Perspectivas futuras**

Embora os esforços envidados para o desenvolvimento da área de avaliação psicológica tenham produzido resultados importantes, vários aspectos ainda podem e devem ser desenvolvidos no Brasil,

como o fortalecimento do compromisso das práticas de avaliação com a inclusão social e o respeito aos direitos humanos, a utilização de recursos informáticos, especialmente para o desenvolvimento da testagem adaptativa computadorizada, emprego de análises estatísticas robustas para análise de grandes quantidades de dados (*big data*), entre outros. No entanto, esses aspectos estão em desenvolvimento em todo o mundo e não são especificidades da realidade brasileira. Dois aspectos, no entanto, merecem atenção prioritária na realidade brasileira: o aprimoramento do Satepsi e a melhoria da formação em avaliação psicológica.

Uma das críticas dirigidas ao Satepsi, por exemplo, é de que os critérios mínimos adotados como referência para avaliação dos instrumentos são pouco exigentes (Anache, & Corrêa, 2010; Primi, & Nunes, 2010). Uma das consequências disso é que não importa o número nem a qualidade dos estudos apresentados nos manuais dos instrumentos, se eles atenderem os critérios mínimos, serão classificados igualmente como “favoráveis” ao uso profissional. Essa prática pode não estar contribuindo para a melhoria da qualidade dos instrumentos. A proposta, nesse caso, seria ampliar o leque de denominações para instrumentos favoráveis, considerando o número e a diversidade dos estudos de fidedignidade, padronização e, especialmente, validade apresentados no manual, para classificá-los em nível de excelência, adequado ou mínimo, por exemplo.

Além disso, é muito mais frequente a apresentação de estudos de validade com base na estrutura interna, especialmente de validade fatorial, do que estudos de validade de critério, que fornecem uma descrição de como o construto de interesse costuma se relacionar com outras variáveis. Esse tipo de estudo é bastante relevante para sustentar resultados de avaliações que afetam significativamente a vida das pessoas avaliadas, como processos seletivos, avaliações educacionais, avaliações solicitadas pelo sistema judiciário, entre outras. Nessas situações as decisões devem ser tomadas com base no processo de avaliação, fundamentado numa proposta teórica e em evidências empíricas que as sustentem. Essas alterações não foram contempladas na recente Resolução do CFP nº 09/2018, mas são caminhos pelos quais o sistema poderia avançar, estimulando a melhoria da qualidade dos instrumentos e, por consequência, fortalece o compromisso da Psicologia com a sociedade.

Outra limitação do Satepsi é que ele atinge principalmente qualidade psicométrica dos instrumentos de medida e a organização dessas informações num manual coerente e útil para os usuários. No entanto, ele atinge apenas indiretamente o problema da qualidade da formação em avaliação psicológica, uma das principais preocupações da área desde há muito tempo e que ainda permanece (Noronha, & Reppold, 2010; Pereira, & Carellos, 1995). De fato, a falha na formação implica em problemas graves e frequentes na opinião de profissionais que atuam com avaliação psicológica, como o mau uso e a má qualidade dos instrumentos, problemas éticos e epistemológicos (Noronha, 2002). Os próprios alunos também consideraram sua formação em avaliação psicológica como insuficiente e boa parte deles sequer conhecia o Sistema de Avaliação de Testes Psicológicos do CFP (Paula, Pereira, & Nascimento, 2007).

Esses estudos sugerem a necessidade de aprimoramento do ensino de avaliação psicológica nas universidades brasileiras. Alguns estudos têm identificado os problemas que dificultam a melhoria da qualidade da formação: 1) a permanência de uma visão limitada e preconceituosa em relação à avaliação psicológica (Alchieri, & Bandeira, 2002; Hutz, & Bandeira, 2003); 2) a falta de qualificação dos professores (Noronha, Barros, Nunes, & Santos, 2014); 3) a carga horária reduzida das disciplinas de avaliação psicológica (Noronha, & Reppold, 2010; Reppold, & Serafini, 2010) e 4) o ensino descontextualizado e apenas procedimentais de testes psicológicos (Alchieri, & Bandeira, 2002).

Iniciativas na direção de mudar esse cenário têm sido realizadas, mas ainda de forma tímida, individual e pouco divulgadas. Um exemplo de iniciativas desse tipo é a do 8º Congresso Brasileiro de Avaliação Psicológica (Florianópolis, 2017), cujo tema foi “A formação em avaliação psicológica e a avaliação psicológica na formação” (<http://www.ibapnet.org.br/congresso2017/>), que ofereceu cursos, mesas redondas e debates sobre o assunto, o que já tem ocorrido em outras edições desse evento, mas nenhuma iniciativa mais concreta de articulação entre os professores para realização de pesquisas sobre o assunto ou mesmo troca de experiências de sala de aula. Nunes et al. (2012), por sua vez, apresentaram uma proposta de organização curricular na área de avaliação psicológica, na qual identificam as competências que as disciplinas de avaliação psicológica devem desenvolver,

sugerem um conjunto de disciplinas para composição do currículo de cursos de Psicologia, apresentam os elementos considerados importantes para o bom desenvolvimento do ensino de avaliação psicológica, como infraestrutura, métodos de ensino, formação docente, entre outras, e sugerem um rol de referências bibliográficas para esse tema.

Talvez esse processo de melhoria da qualidade da formação em avaliação psicológica tenha que ser mais longo do que se gostaria, pois envolve questões político-administrativas como a necessidade de espaço, tanto físico quanto na grade horária, e estrutura compatíveis com a formação que se deseja. Uma ação política que poderia ajudar a melhorar a formação, que tem sido reivindicada pelos profissionais da área, é o reconhecimento da especialidade em avaliação psicológica pelo CFP. O CFP já reconhece o especialista em diversas áreas, como clínica, escolar, hospitalar, organizacional, entre outras, mas não em avaliação psicológica. Um documento em favor desse reconhecimento, assinado pelo IBAP, pela ASBRO e pela Comissão Consultiva em Avaliação Psicológica (gestão de 2012 a 2014), apontava as seguintes vantagens desse reconhecimento: 1) valorização das disciplinas de avaliação psicológica no currículo da graduação; 2) fomento da criação de cursos de formação especializada; 3) estímulo, entre os profissionais, da formação continuada, valorizando aqueles que buscarem qualificação nesse campo e promovendo a carreira desses

profissionais; 4) reconhecimento do mérito de quem já se dedica a esta área; e 5) redução dos problemas de má conduta ética, beneficiando a sociedade com a prestação de serviços de qualidade. Mesmo assim, o conselho vem rejeitando essa proposta, alegando que a avaliação psicológica é uma área de formação básica e comum a todos os psicólogos e não uma área de especialidade. Essa alegação contrasta com todas as outras ações de valorização da área de avaliação psicológica promovidas pelo CFP nas últimas décadas, dificultando, ou deixando de incentivar e reconhecer a formação continuada nessa área.

### Considerações finais

Como se pode perceber, a área de avaliação psicológica passou por grandes mudanças no Brasil, que, em grande parte, refletem os movimentos observados em outras partes do mundo. Tanto as críticas que recebeu, e continua recebendo, quanto as ações que foram implementadas para fortalecê-la, contribuíram significativamente para o seu desenvolvimento. Consequentemente a própria Psicologia também se desenvolveu, pois bons instrumentos de avaliação permitem investigações (produção do conhecimento) e favorecem a atuação de psicólogos em todas as áreas da atividade humana. Assim, a avaliação psicológica contribuiu e continua contribuindo para o estabelecimento da Psicologia como ciência e profissão também no cenário brasileiro.

### Referências

- Alchieri, J. C., & Bandeira, D. R. (2002). O ensino da avaliação psicológica no Brasil. In R. Primi (Ed.), *Temas em avaliação psicológica* (pp. 35-39). São Paulo, SP: Instituto Brasileiro de Avaliação Psicológica.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Anache, A., & Corrêa, F. (2010). As políticas do conselho federal de psicologia para a avaliação psicológica. In Conselho Federal de Psicologia – CFP (Ed.), *Avaliação psicológica: Diretrizes na regulamentação da profissão* (pp. 19-30). Recuperado de <https://site.cfp.org.br/publicacao/avaliacao-psicologica-diretrizes-na-regulamentacao-da-profissao/>
- Barroso, S. M. (2010). Avaliação psicológica: Análise das publicações disponíveis na SciELO e BVS-Psi. *Fractal: Revista de Psicologia*, 22(1), 141-154. <https://doi.org/10.1590/S1984-02922010000100011>
- Binet, A. (1904). A propos de la mesure de l'intelligence. *L'Année Psychologique*, 11(1), 69-82. <https://doi.org/10.1037/h0068738>
- Bueno, J. M. H., & Ricarte, M. D. (2017). Aspectos históricos da testagem psicológica: contexto internacional e nacional. In: M. R. C. Lins, & J. C. Borsa (Orgs.), *Avaliação psicológica: Aspectos teóricos e práticos* (pp. 38-55). Petrópolis, RJ: Vozes.

- Bueno, J. M. H., Amorim, D. A., & Albuquerque, E. S. G. (2017). Implantação do SATEPSI e impactos na área de avaliação psicológica. In: B. F. Damásio, J. C. Borsa (Orgs.), *Manual de desenvolvimento de instrumentos psicológicos* (pp. 119-137). São Paulo, SP: Vetor.
- Cattell, J. M. (1890). Mental tests and measurements. *Mind*, 15, 373-380.
- Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika*, 12, 197-220. <https://doi.org/10.1007/BF02289253>
- Cattell, R. B. (1946). *The description and measurement of personality*. New York, NY: World Book.
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40, 153-193. <https://doi.org/10.1037/h0059973>
- Gouveia, V. V. (2009). A avaliação psicológica no Brasil: caminhos, desafios e possibilidades. *Psicologia em Foco*, 2(1), 110-119.
- Hutz, C. S. (2015). Questões éticas na avaliação psicológica. In: C. S. Hutz, D. R. Bandeira, & C. M. Trentini (Orgs.), *Psicométrica* (pp. 165-173). Porto Alegre, RS: Artmed.
- Hutz, C. S. (2002). Responsabilidade ética, social e política da avaliação psicológica. *Avaliação Psicológica*, 1(2), vii-ix.
- Hutz, C. S., & Bandeira, D. S. (2003). Avaliação psicológica no Brasil: Situação atual e desafios para o futuro. In O. H. Yamamoto, & V. V. Gouveia (Eds.), *Construindo a psicologia brasileira: Desafios para a ciência e prática psicológica* (pp. 261-277). São Paulo, SP: Casa do Psicólogo.
- Joly, M. C. R. A., Silva, M. C. R. D., Nunes, M. F. O., & Souza, M. S. D. (2007). Análise da produção científica em painéis dos congressos brasileiros de avaliação psicológica. *Avaliação Psicológica*, 6(2), 239-252.
- Jorm, A. F. (2015). Using the Delphi expert consensus method in mental health research. *Australian & New Zealand Journal of Psychiatry*, 49(10), 887-897. <https://doi.org/10.1177/0004867415600891>
- Koocher, G. P., & Keith-Spiegel, P. (2016). Psychological assessment: Testing tribulations. In: G. P. Koocher, & P. Keith-Spiegel (Eds.), *Ethics in psychology and the mental health professions: Standards and cases* (pp. 193-232). New York, NY: Oxford University Press.
- Laher, S., & Cockcroft, K. (2016). Contextualizing psychological assessment in South Africa. In: S. Laher, & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and Applications* (pp. 1-14). Johannesburg: Wits University Press.
- Lei Nº 4.119, de 27 de agosto de 1962. Dispõe sobre os cursos de formação em psicologia e regulamenta a profissão de psicólogo. *Diário Oficial da União*, 5 set. 1962.
- Leong, F. T. L., & Park, Y. S. (2016). Introduction. In: Council of National Psychological Associations for the Advancement of Ethnic Minority Interests – CNPAAEMI (Org.) *Testing and assessment with persons & communities of color* (pp. 1-2). Washington, DC: o autor.
- Mansur-Alves, M., Silva, R. S., & Fernandes, S. C. A. (2016). Impact of the Psychological Testing Assessment System (SATEPSI) for scientific publications in psychological assessment. *Psico-USF*, 21(1), 179-188. <https://doi.org/10.1590/1413-82712016210115>
- Messick, S. (1989). Meaning and values in test validation: the science and ethics of assessment. *Educational Researcher*, 18(2), 5-11.
- Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L. et al. (2006). Protocol of the COSMIN study: CONsensus-based Standards for the selection of health Measurement INSTRUMENTS. *BMC Medical Research Methodology*, 6(2), (open access). <https://doi.org/10.1186/1471-2288-6-2>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L. et al. (2012). COSMIN checklist manual. Amsterdam: VU University Medical Centre. Recuperado de <http://www.cosmin.nl/index.html>
- Mokkink, L. B.; Terwee, C. B.; Patrick, D. L.; Alonso, J.; Stratford, P. W.; Knol, D. L. et al. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, 19(4), 539-549. doi: 10.1007/s11136-010-9606-8

- Noronha, A. P. P. (2002). Os problemas mais graves e mais frequentes no uso dos testes psicológicos. *Psicologia: Reflexão e Crítica*, 15(1), 135-142. <https://doi.org/10.1590/S0102-79722002000100015>.
- Noronha, A. P. P., Barros, M. V. C., Nunes, M. F. O., & Santos, A. A. A. (2014). Avaliação psicológica: importância e domínio de atividades segundo docentes. *Estudos e Pesquisas em Psicologia*, 14(2), 524-538.
- Noronha, A. P. P., Primi, R., & Alchieri, J. C. (2004). Parâmetros psicométricos: uma análise de testes psicológicos comercializados no Brasil. *Psicologia: Ciência e Profissão*, 24(4), 88-89. <https://doi.org/10.1590/S1414-98932004000400011>
- Noronha, A. P. P., & Reppold, C. T. (2010). Considerações sobre a avaliação psicológica no Brasil. *Psicologia Ciência e Profissão*, 30(no. Spe), 192-201. <https://doi.org/10.1590/S1414-98932010000500009>
- Nunes, M. F. O., Muniz, M., Reppold, C. T., Faiad, C., Bueno, J. M. H., & Noronha, A. P. P. (2012). Diretrizes para o ensino de avaliação psicológica. *Avaliação Psicológica*, 11, 309-316.
- Ocampo, M. L. S. D., Arzeno, M. E. G., & Piccolo, E. D. (1981). *O processo psicodiagnóstico e as técnicas projetivas*. São Paulo, SP: Martins Fontes.
- Pasquali, L., & Alchieri, J. C. (2011). Os testes psicológicos no Brasil. In: L. Pasquali (Org.), *Técnicas de Exame Psicológico – TEP: Manual* (2a ed., pp. 195-221). São Paulo, SP: Casa do Psicólogo..
- Patto, M. H. S. (1997). Para uma crítica da razão psicométrica. *Psicologia USP*, 8(1), 47-62. <https://doi.org/10.1590/S0103-65641997000100004>
- Paula, A. V.; Pereira, A. S.; Nascimento, E. (2007). Opinião de alunos de psicologia sobre o ensino em avaliação psicológica. *PsicoUSF*, 12(1), 33-43.
- Pereira, A. P. C., & Carellos, S. D. M. S. (1995). Examinando o ensino das técnicas de exame psicológico. *Caderno de Psicologia*, 3(4), 33-36.
- Primi, R. (2010). Avaliação psicológica no Brasil: fundamentos, situação atual e direções para o futuro. *Psicologia: Teoria e Pesquisa*, 26(n.spe), 25-35. <https://doi.org/10.1590/S0102-37722010000500003>
- Primi, R., Muniz, M., & Nunes, C. H. N. S. (2009). Definições contemporâneas de validade de testes psicológicos. In: C. S. Hutz (Org.), *Avanços e polêmicas em avaliação psicológica* (pp. 243-265). São Paulo, SP: Casa do Psicólogo
- Primi, R., & Nunes, C. H. S. (2010). O Satepsi: Desafios e propostas de aprimoramento. In: Conselho Federal de Psicologia – CFP (Ed.), *Avaliação psicológica: Diretrizes na regulamentação da profissão* (pp. 129-148). Brasília, DF: o autor.
- Reppold, C. T., & Serafini, A. J. (2010). Novas tendências no ensino da avaliação psicológica. *Avaliação Psicológica*, 9(2), 323-329.
- Resolução Nº 002, de 24 de março de 2003*. Define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos e revoga a Resolução CFP nº 025/2001. Brasília, DF: Conselho Federal de Psicologia.
- Resolução Nº 009, de 25 de abril de 2018*. Estabelece diretrizes para a realização de Avaliação Psicológica no exercício profissional da psicóloga e do psicólogo, regulamenta o Sistema de Avaliação de Testes Psicológicos - SATEPSI e revoga as Resoluções nº 002/2003, nº 006/2004 e nº 005/2012 e Notas Técnicas nº 01/2017 e 02/2017. Brasília, DF: Conselho Federal de Psicologia.
- Resolução Nº 25, de 30 de novembro de 2001*. Define teste psicológico como método de avaliação privativo do psicólogo e regulamenta sua elaboração, comercialização e uso. Brasília, DF: Conselho Federal de Psicologia.
- Spearman, C. (1904). "General intelligence", objectively determined and measured. *The American Journal of Psychology*, 15, 201-293.
- Sue, D. W., Arredondo, P., & McDavis, R. J. (1992). Multicultural counseling competencies and standards: A call to the profession. *Journal of Counseling & Development*, 70, 477-486. <https://doi.org/10.1002/j.1556-6676.1992.tb01642.x>
- Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research*, 21(4), 651-657. <https://doi.org/10.1007/s11136-011-9960-1>
- Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs*, 1.

Thurstone, L. L. (1937). Psychology as a quantitative rational science. *Science*, 85, 227-232.

Thurstone, L. L. (1931). *The reliability and validity of tests*. Ann Arbor, MI: Edward Brothers.

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, 41, 1-32. <https://doi.org/10.1037/h0075959>

Zaia, P., Oliveira, K. S., & Nakano, T. C. (2018). Análise dos processos éticos publicados no Jornal do Conselho Federal de Psicologia. *Psicologia: Ciência e Profissão*, 38(1), 8-21. <https://doi.org/10.1590/1982-3703003532016>

---

*José Maurício Haas Bueno*

Doutor em Psicologia (2008) pelo Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco. É professor da Universidade Federal de Pernambuco, no Curso de Graduação em Psicologia e no Programa de Pós-Graduação em Psicologia Cognitiva. Realiza, orienta e publica trabalhos de investigação científica, com ênfase na construção, adaptação e investigação das propriedades psicométricas de instrumentos de avaliação psicológica.

E-mail: mauricio.ufpe@gmail.com

*Evandro Moraes Peixoto*

Docente do Departamento de Psicologia da Universidade de Pernambuco - UPE. Coordenador do Laboratório de Avaliação Psicológica e Psicometria LAPPsi. Pós-doutorando em Psicologia pela Universidade São Francisco-USE. Doutor em Psicologia como Profissão e Ciência pela Pontifícia Universidade Católica de Campinas PUCC, com estágio doutoral PDSE desenvolvido na Université du Québec à Trois-Rivières - QC Canadá. Membro do grupo de trabalho Avaliação Psicológica em Psicologia Positiva e Criatividade na ANPEPP.

E-mail: evandro.peixoto@upe.br

Endereço para envio de correspondência:

Maurício Bueno – Universidade Federal de Pernambuco (UFPE), Centro de Filosofia e Ciências Humanas (CFCH), PPG Psicologia Cognitiva

Av. da Arquitetura, s/n, 8º Andar, CEP: 50740-550.

Recife – PE. Brasil.

*Recebido:* 21/07/2018

*Aprovado:* 08/08/2018

*Received:* 07/21/2018

*Approved:* 08/08/2018

*Recibido:* 21/07/2018

*Aceptado:* 08/08/2018

*Como citar:* Bueno, J. M. H., & Peixoto, E. M. (2018). Avaliação Psicológica no Brasil e no Mundo. *Psicologia: Ciência e Profissão*, 38(n.spe), 108-121. <https://doi.org/10.1590/1982-3703000208878>

*How to cite:* Bueno, J. M. H., & Peixoto, E. M. (2018). Psychological Assessment in Brazil and in the World. *Psicologia: Ciência e Profissão*, 38(n.spe), 108-121. <https://doi.org/10.1590/1982-3703000208878>

*Cómo citar:* Bueno, J. M. H., & Peixoto, E. M. (2018). Evaluación Psicológica en Brasil y en el Mundo. *Psicologia: Ciência e Profissão*, 38(n.spe), 108-121. <https://doi.org/10.1590/1982-3703000208878>