

**UNIVERSIDADE FEDERAL DO MARANHÃO**  
**CENTRO DE CIÊNCIAS HUMANAS**  
**DEPARTAMENTO DE PSICOLOGIA**

**DAPHINE ROLIM VELOSO**

**REVISÃO INTEGRATIVA SOBRE AS PROPRIEDADES PSICOMÉTRICAS**  
**DE INSTRUMENTOS DESENVOLVIDOS NO BRASIL**

**São Luís**

**2018**

**DAPHINE ROLIM VELOSO**

**REVISÃO INTEGRATIVA SOBRE AS PROPRIEDADES PSICOMÉTRICAS  
DE INSTRUMENTOS DESENVOLVIDOS NO BRASIL**

Monografia apresentada ao Departamento de Psicologia da Universidade Federal do Maranhão como requisito para obtenção do título de Psicólogo  
Orientador: Prof. Dr. Lucas Guimarães Cardoso de Sá  
Área de concentração: Psicometria e Avaliação Psicológica

**São Luís**

**2018**

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).  
Núcleo Integrado de Bibliotecas/UFMA

Veloso, Daphine Rolim.

Revisão integrativa sobre as propriedades psicométricas  
de instrumentos desenvolvidos no Brasil / Daphine Rolim  
Veloso. - 2018.

93 f.

Orientador(a): Lucas Guimarães Cardoso de Sá.

Monografia (Graduação) - Curso de Psicologia,  
Universidade Federal do Maranhão, São Luís, 2018.

1. Avaliação Psicológica. 2. Normas. 3. Precisão. 4.  
Psicometria. 5. Validade. I. Sá, Lucas Guimarães Cardoso  
de. II. Título.

**DAPHINE ROLIM VELOSO**

**REVISÃO INTEGRATIVA SOBRE AS PROPRIEDADES PSICOMÉTRICAS  
DE INSTRUMENTOS DESENVOLVIDOS NO BRASIL**

Monografia apresentada ao Departamento de Psicologia da Universidade Federal do Maranhão como requisito para obtenção do título de Psicólogo.

Aprovada em \_\_\_/\_\_\_/\_\_\_.

---

Prof. Dr. Lucas Guimarães Cardoso de Sá (Orientador)  
Doutorado em Psicologia (Universidade Federal de São Carlos)  
Universidade Federal do Maranhão

---

Prof. Dra. Ana Beatriz Rocha Lima  
Doutorado em Psicologia Social, do Trabalho e das Organizações (Universidade de Brasília)  
Universidade CEUMA

---

Prof. Dr. Tony Nelson  
Doutorado em Psicologia (Universidade Federal do Pará)  
Universidade Federal do Maranhão

**São Luís**

**2018**

## AGRADECIMENTOS

Antes de tudo a minha gratidão a Deus, sem Ele, todos os meus planos seriam frustrados;

À minha família, em especial, aos meus pais (Damiana e Davilson), minha irmã (Stephane), meu cunhado (Atila), meus avós (Assis e Lucinete) e minhas primas (Luiza e Mariana) que nunca, nem por um segundo, soltaram minha mão, sempre fazendo o possível para que eu pudesse seguir. Sou muito agraciada por ter cada um de vocês na minha vida. A melhor família que há.

À Erica Mayara, o maior presente que ganhei da Psicologia, quem sinto do meu lado, de muitas maneiras, apesar da enorme distância física; à Aby, que me acompanhou de perto em vários momentos dessa jornada, obrigada pelo apoio, minha querida amiga, que tudo te vá bem; a Emanuel, por tornar meus dias mais calmos. Não tem nome pro tanto de você que mora em mim.

Às pessoas que conheci na UFMA, levo um pouquinho de vocês comigo. A existência de vocês fez com que o caminho fosse menos árduo.

Ao professor Lucas, um dos maiores exemplos de dedicação que tenho... uma das pessoas que me mostrou como a Psicologia pode ser bonita; o seu trabalho foi e é um dos pontos que mais pesaram na minha permanência e profundo encantamento pela Psicologia. Sem medo de errar, os dados atuais (ausência de artigos na temática, publicados por pessoas vinculadas ao Maranhão), estão mudando, em grande por sua presença. Minha admiração;

Ao professor Tony, que desde o projeto, fez apontamentos pertinentes e indispensáveis;

À minha supervisora de estágio, Edla Maria e à Mary (Assistente Social), o exemplo de vocês me captura e às minhas colegas de estágio que se tornaram boas amigas.

Aos componentes da banca, Ana Beatriz, Tony Nelson e Yuri, que de pronto aceitaram o convite para contribuir com o aprimoramento deste trabalho. A “boa”

ciência é feita de pessoas como vocês. Obrigada por partilharem conhecimento e contribuírem para que esse trabalho fosse aprimorado.

Enfim, a todos os meus amigos e companheiros de jornada, sou um pouquinho de cada um.

## RESUMO

O processo de Avaliação Psicológica é uma atividade do psicólogo cada vez mais demandada pela sociedade. Assim, é esperado que exista uma preocupação constante com a qualidade das diversas técnicas utilizadas na sua realização. Considerando que a testagem psicológica é, provavelmente, a técnica mais utilizada nesse processo, a ponto de ser confundida com a própria avaliação, este estudo tem por objetivo investigar as propriedades psicométricas dos instrumentos desenvolvidos no Brasil, no período entre 2003 e 2017. O método escolhido foi o de revisão integrativa, em que foram selecionados os resumos de artigos identificados na base de dados *Scielo*, a partir de seis palavras mais frequentes que foram sugeridas por seis especialistas na área: validade, precisão, fidedignidade, confiabilidade, normas e psicometria. Os resultados indicaram que a fonte de validade mais comum é a baseada em variáveis externas; no que diz respeito à precisão, a mais investigada é a consistência interna; já em normas, somente três dos estudos especificaram nos resumos quais os tipos foram utilizados. Além disso, identificou-se que a região Norte do Brasil é a que tem a menor quantidade de pesquisas no foco desta pesquisa, enquanto os maiores números foram encontrados no eixo Sul-Sudeste. Em relação ao Nordeste, a Paraíba lidera os estudos sobre características psicométricas e o Maranhão é o único estado da região a não ter nenhuma pesquisa na temática. Conclui-se que, apesar do crescimento nos estudos das características psicométricas dos instrumentos desenvolvidos no Brasil, muitas mudanças ainda são necessárias, tanto no que diz respeito à melhor formação quanto na integração entre instituições, a fim de que os estudos avancem sempre na direção de instrumentos de qualidade.

Palavras - chave: Psicometria, Validade, Precisão, Normas, Avaliação psicológica.

## ABSTRACT

The process of Psychological Assessment is an activity of the psychologist increasingly demanded by society. Thus, it is expected that there is a constant concern with the quality of the various techniques used in its accomplishment. Considering that psychological testing is probably the most used technique in this process, to the point of being confused with the evaluation itself, this study aims to investigate the psychometric properties of the instruments developed in Brazil from 2003 to 2017. The chosen method was the selection of the abstracts of articles identified in the *Scielo* database, with more frequent words that were suggested by six specialists in the area: Validity, precision, reliability, norms and psychometrics. The results indicated that the most common source of validity is based on external variables; with regard to reliability, the most investigated is the internal consistency; in norms, only three of the studies specified in the abstracts which types were used. Besides, it was identified that the North region is the one that has the least amount of research in the focus of this research, while the largest numbers were found in the South-Southeast axis. In relation to the Northeast, Paraíba State leads the studies investigating psychometric characteristics and Maranhão is the only State in the region to have no research on the subject. It is concluded that, despite the growth in the studies of psychometric characteristics, many changes are still necessary, both in terms of better training and in the integration between institutions, so that studies always move towards quality instruments.

Key words: Psychometrics, Validity, Reliability, Norms, Psychological assessment.



## **LISTA DE GRÁFICOS**

<b>GRÁFICO 1 -</b>	<b>Comparativo entre as principais fontes de validade</b>	<b>30</b>
<b>GRÁFICO 2 -</b>	<b>Comparativo entre as principais fontes de precisão</b>	<b>34</b>

## **LISTA DE TABELAS**

<b>TABELA 1 -</b>	<b>Palavras-chave sugeridas pelos juízes especialistas</b>	<b>27</b>
<b>TABELA 2 -</b>	<b>Frequência das palavras por semelhança</b>	<b>28</b>
<b>TABELA 3 -</b>	<b>Quantidade de artigos por categorias</b>	<b>29</b>
<b>TABELA 4 -</b>	<b>Produção por Regiões do Brasil (Ordem Decrescente)</b>	<b>37</b>
<b>TABELA 5 -</b>	<b>Produção por Estados do Nordeste</b>	<b>38</b>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	9
1.1	JUSTIFICATIVA	10
1.2	PROBLEMA DE PESQUISA	11
1.3	OBJETIVOS	11
1.3.1	Geral	11
1.3.2	Específicos	11
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	12
2.1	CONTEXTUALIZAÇÃO	12
2.2	VALIDADE	16
2.3	PRECISÃO	18
2.4	NORMAS	21
<b>3</b>	<b>MÉTODO</b>	24
3.2	BUSCA NA LITERATURA E CRITÉRIOS DE SELEÇÃO DOS ARTIGOS	24
3.3	APRESENTAÇÃO DOS RESULTADOS	25
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	26
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	40
	<b>REFERÊNCIAS</b>	42
	<b>APÊNDICES</b>	45
	<b>APÊNDICE A - Tabela das fontes de validade referentes à palavra-chave Psicometria</b>	46
	<b>APÊNDICE B - Tabela das fontes de validade referentes à palavra-chave validade</b>	47
	<b>APÊNDICE C - Tabela das fontes de validade referentes à palavra-chave fidedignidade</b>	48
	<b>APÊNDICE D - Tabela das fontes de validade referentes à palavra-chave precisão</b>	49
	<b>APÊNDICE E - Tabela das fontes de validade referentes à palavra-chave confiabilidade</b>	50
	<b>APÊNDICE F - Tabela dos tipos de precisão referentes à palavra-chave Psicometria</b>	51
	<b>APÊNDICE G - Tabela dos tipos de precisão referentes à palavra-chave validade</b>	52
	<b>APÊNDICE H - Tabela dos tipos de precisão referentes à palavra-chave fidedignidade</b>	53
	<b>APÊNDICE I - Tabela dos tipos de precisão referentes à palavra-chave precisão</b>	54
	<b>APÊNDICE J - Tabela dos tipos de precisão referentes à palavra-chave confiabilidade</b>	55
	<b>APÊNDICE K - Relação de artigos: palavra-chave psicometria</b>	56
	<b>APÊNDICE L - Relação de artigos: palavra-chave validade</b>	65
	<b>APÊNDICE M - Relação de artigos: palavra-chave fidedignidade</b>	87
	<b>APÊNDICE N - Relação de artigos: palavra-chave precisão</b>	89
	<b>APÊNDICE O - Relação de artigos: palavra-chave normas</b>	91
	<b>APÊNDICE P - Relação de artigos: palavra-chave confiabilidade</b>	93

## 1 INTRODUÇÃO

A testagem e avaliação psicológicas passam por um momento tanto de discussões acerca de questões éticas, quanto de maior presença nas diversas áreas de atuação. Tudo isso é fruto do percurso histórico que as acompanha. Cabe destacar, já que, não raro, os termos são utilizados como sinônimos, que as regulamentações dos últimos 15 anos focavam essencialmente na testagem psicológica e não na avaliação psicológica. Entretanto, a Resolução CFP 009/2018, de abril de 2018, explicita com maior amplitude a importância da Avaliação Psicológica.

A testagem tem por finalidade obter uma medida de um atributo ou característica por meio de um instrumento, conhecido usualmente como teste (COHEN; SWERDLIK; STURMAN, 2014). De modo geral, ele é uma ferramenta que avalia os resultados com base em padrões de comparação (URBINA, 2007). Por sua vez, a avaliação psicológica diz respeito a um processo, com integração de diferentes técnicas, como entrevistas, estudos de caso, observação, técnicas de dinâmica de grupo, análise de documentos e testagem, com o objetivo de coletar informações, estabelecer e confirmar hipóteses que embasem a tomada de decisões (COHEN; SWERDLIK; STURMAN, 2014). Ela pode ser utilizada com diferentes objetivos (por exemplo, descrição, diagnóstico, predição, monitoramento de comportamento), em diversos contextos, dentre os quais se destacam: porte de armas, concursos públicos, sistemas judiciário e prisional, clínica, trânsito, cirurgia bariátrica, recursos humanos, orientação profissional e na área psicoeducacional, entre outras (CONSELHO FEDERAL DE PSICOLOGIA, 2013; MACHADO; MORONA, 2007).

Justamente pelo peso social que a tomada de decisões a partir de uma avaliação psicológica possui, o desconhecimento do processo e os processos éticos resultantes são motivos que devem ser pesados ante a importância de estudos nessas áreas, que são privativas do psicólogo. Considerando a testagem psicológica como parte fundamental do processo, o estabelecimento de critérios mínimos para a utilização de testes na prática profissional é indispensável para a melhoria na qualidade dos instrumentos que são utilizados e para o consequente aumento da credibilidade e da valorização profissional.

Este trabalho é apresentado da seguinte forma: primeiramente uma fundamentação teórica contendo a contextualização da temática e explanação dos critérios mínimos preconizados pelo Sistema de Avaliação de Teste Psicológicos

(SATEPSI), com seus devidos métodos de investigação. Em seguida, apresentação do método escolhido para a pesquisa e posterior apresentação e discussão dos resultados, com considerações finais para fechar o trabalho.

## 1.1 JUSTIFICATIVA

A percepção de que os testes disponíveis no Brasil possuem problemas parece unânime (PRIMI; NUNES, 2010). Diante disso, surge a necessidade de investigar como os psicometristas tem tratado questões de validade, precisão e normas dos instrumentos desenvolvidos, pois são esses instrumentos que poderão ser comercializados e utilizados nas práticas profissionais dos psicólogos.

Os mesmos autores, Primi e Nunes (2010), ao discorrerem sobre os requisitos mínimos de qualidade de um teste, tecem a crítica de que, atualmente, os critérios de fato são muito básicos, o que ocasiona na aprovação idêntica de testes bem complexos e outros que tangenciam muito de perto o limite mínimo estabelecido. Consideram ainda outro aspecto como relevante: a questão do contexto ao qual o teste é aplicado e se existem estudos com grupo normativo para os grupos específicos que deverão se beneficiar da aplicação do instrumento.

Esses aspectos têm gerado discussões e empenho sobre a necessidade de atenção e até elevação dos critérios mínimos para que um teste seja considerado apto para uso profissional. Sobre isso, Gouveia (2009) enfatiza que não se pode pensar de forma amadora, pois a vida de muitas pessoas poderá ser mudada por meio das consequências geradas pela decisão tomada decorrentes do processo de Avaliação Psicológica.

Para Urbina (2007) diferentes são as finalidades possíveis de escores de um teste, assim, as bases de sua interpretação podem ser provenientes de diferentes métodos:

As contribuições para evidências de validade de escores podem ser feitas por qualquer pesquisa sistemática que corrobore ou acrescente algo ao seu sentido, independente de quem a conduz ou quando ela ocorre. Desde que existam evidências científicas sólidas para um uso proposto dos escores de um teste [...] Esta proposição ajuda a explicar a natureza multifacetada da pesquisa de validação, bem como seus achados muitas vezes redundantes e, às vezes, conflitantes. Também explica a longevidade de alguns instrumentos, como o MMPI e as escalas Wechsler, sobre os quais foi acumulada uma vasta literatura – que engloba numerosas aplicações em uma variedade de contextos – ao longo de décadas de pesquisa (URBINA, p. 156, 2007).

Desse modo, discorrer sobre as qualidades psicométricas de um teste é de grande relevância para a continuidade dos avanços na área, já que cumpre o papel de investigação acerca da qualidade da produção do material envolvido no processo de Avaliação Psicológica e do respeito aos princípios éticos. Como consequência, poderá informar à comunidade científica sobre a importância de se ter dados sobre a qualidade dos testes, desmistificar conceitos pré-estabelecidos, a exemplo da ideia de que os testes são mecanicistas e não avaliam o que se propõem, além de trazer dados sobre as práticas em Psicometria utilizados em território nacional. Por meio de uma Revisão Integrativa (RI), um trabalho de mapeamento dessas práticas no momento atual poderá indicar caminhos a serem seguidos em uma nova etapa histórica da avaliação psicológica no Brasil, que começa agora e se estenderá aos próximos anos.

## 1.2 PROBLEMA DE PESQUISA

A questão norteadora do estudo de revisão integrativa foi: “Quais as propriedades psicométricas utilizadas nos instrumentos desenvolvidos no Brasil?”. A partir dessa questão foram estabelecidas outras mais específicas: (a) Quais as fontes de evidências de validade mais utilizadas? (b) Quais os métodos de precisão mais frequentes? (c) Quais os tipos de normas mais comuns?

## 1.3 OBJETIVOS

### 1.3.1 Geral

Realizar uma revisão sobre as principais técnicas psicométricas utilizadas em pesquisa no Brasil.

### 1.3.2 Específicos

- Identificar as fontes de evidências de validade mais investigadas;
- Analisar métodos de precisão mais utilizados;
- Caracterizar tipos de normas mais comuns nos instrumentos desenvolvidos no Brasil;
- Comparar a frequência das técnicas de acordo com o período (tempo).

## 2 REFERENCIAL TEÓRICO

Considerar historicamente as bases conceituais e práticas de uma área leva à melhor compreensão do seu contexto atual. Assim, não apenas descrever sobre os fundamentos e critérios adotados na construção e apreciação dos testes psicológicos, mas também caracterizar o contexto em que surgiu a Avaliação Psicológica no Brasil e os órgãos regulamentadores faz-se necessário nesse estudo.

### 2.1 CONTEXTUALIZAÇÃO

De acordo com Anastasi e Urbina (2000), as raízes da testagem estão perdidas na antiguidade. Há relatos da utilização de exames no serviço civil no império chinês, há dois mil anos. Entre os gregos, testes eram feitos no processo educacional e, na Idade Média, utilizavam-se testes para conferir graus e títulos, nas universidades europeias. No século XIX, destacou-se o trabalho do biólogo inglês Francis Galton, que foi o pioneiro na utilização de métodos estatísticos, questionários e avaliação. Por volta do ano de 1884 esse cientista já realizava exames sensoriais, perceptuais e de medidas físicas nos laboratórios antropométricos e acreditava que tais características estavam relacionadas à inteligência. Suas ideias baseavam-se na hereditariedade e geraram críticas por darem argumentos para a eugenia. Seu trabalho influenciou Charles Spearman e Karl Pearson e possibilitou a criação do conceito de correlação estatística entre variáveis (ANASTASI; URBINA, 2000).

Embora os trabalhos de Galton devam ser mencionados como um aspecto histórico da testagem psicológica, Cohen, Swerdlik e Sturman (2014); Anastasi e Urbina (2000) indicam que a França é o berço da testagem e avaliação psicológicas modernas, tendo como marco principal o desenvolvimento, em 1905, do teste elaborado por Alfred Binet (considerado o inventor dos testes de inteligência) e seu colega, Theodore Simon (psicólogo francês, que fez estudos sobre pessoas com deficiência mental). Esse instrumento averiguava a “Idade Mental” de estudantes, a partir da capacidade atual desses indivíduos em comparação com outros da mesma idade e tinha como finalidade colocar as crianças em classes de aula adequadas nas escolas parisienses. O instrumento se disseminou e, dentro de uma década, estava nas instituições escolares dos Estados Unidos, onde se destaca o nome de Lewis Terman como principal pesquisador que trabalhou com o instrumento desenvolvido por Binet, a

ponto de praticamente criar um novo instrumento baseado neste, a escala Stanford-Binet.

No período da Primeira Guerra Mundial, por volta de 1917, com a entrada dos EUA no conflito, criou-se a necessidade de avaliar de maneira rápida as características emocionais e intelectuais dos recrutas. A testagem, até então de caráter individual, passou a ser desenvolvida e utilizada em proposta de aplicação coletiva com fins de classificação, seleção, admissão e dispensa de militares, o que se repetiu na Grande Segunda Guerra Mundial. Destaca-se nesse processo inicial o nome de Robert Yerkes (COHEN; SWERDLIK; STURMAN, 2014; ANASTASI; URBINA, 2000).

Segundo Vieira e Campos (2011) o pensamento e a cultura no Brasil, em torno do século XIX e início do século XX, foi fortemente influenciado pelas ideias francesas. Diante dessas circunstâncias, as primeiras concepções sobre testes, primordialmente difundidas na área Médica e da Educação, também procederam dos estudos nesse país. Nesse transcurso de tempo, surgiram as primeiras demandas vindas do meio educacional, o que desembocou no maior interesse pelo desenvolvimento de instrumentos de testagem. Dentre os precursores desse momento de pesquisa e criação estão nomes como Ugo Pizzolli e Lourenço Filho, um dos maiores divulgadores da obra de Binet. É importante salientar que a Psicologia não entrou de forma acessível no país, sempre estando envolta em tabus e preconceitos, com uma imagem mística e oculta (VIEIRA; CAMPOS, 2011).

No período entre 1932 a 1962, ocorreu intensa aplicação de testes em indústrias, escolas e clínicas (VIEIRA; CAMPOS, 2011), sendo também dessa época a inserção da avaliação psicológica no contexto do trânsito. Em decorrência do avanço da indústria automobilística e necessidade por segurança, em 1951 a testagem passou a ser requisito para retirar Carteira Nacional de Habilitação (SILVA; ALCHIERI, 2007). Em 1962 ocorreu a regulamentação da Psicologia e nos quinze anos seguintes os cursos se expandiram, porém, ao contrário do que poderia se esperar, a qualidade da formação não acompanhou esse movimento de expansão. Assim, a pesquisa em testagem e avaliação psicológica perdeu seu espaço. Somente a partir de 1980 voltaram, aos poucos, a receber atenção (VIEIRA; CAMPOS, 2011).

No entanto, o período anterior, de explosão dos cursos de Psicologia e pouca qualidade da formação em Avaliação Psicológica obviamente deixou sequelas no que diz respeito à formação e qualificação profissional em relação à Avaliação Psicológica. Conseqüentemente, o descrédito e banalização do uso de testes tornou-se

comum. O uso incorreto, o pouco conhecimento teórico e técnico, a desvalorização, a falta de padronização, avaliações subjetivas e poucos estudos, além de boa parte do material utilizado ter sido mal adaptado e inadequado à realidade nacional foram só alguns dos problemas descritos (GOMES, 2003; VIEIRA; CAMPOS, 2011).

Como forma de reduzir os danos, ao final da década de 1990, o Conselho Federal de Psicologia (CFP) começou a discutir o estado da área no Brasil, o que culminou na oficialização, em 2002, de uma Comissão Consultiva em Avaliação Psicológica, a fim de analisar os testes usados no Brasil (VIEIRA; CAMPOS, 2011; PRIMI, 2010). Outrossim, foi criado, em 2003, o Sistema de Avaliação dos Testes Psicológicos (SATEPSI), com funções de certificação para uso profissional, avaliação e qualificação de instrumentos em aptos e inaptos.

O nascimento do SATEPSI deu-se pela necessidade de regulamentação de um processo avaliativo que mostrasse evidências mínimas de eficiência dos instrumentos, o que em relação à quantidade considerável de processos éticos, não ocorria até então. Inicialmente ele parece não ter sido bem visto pelos profissionais e casasadoras, que acreditavam ser um empecilho na comercialização. No entanto, as consequências, como melhora na qualidade do material produzido, comprovaram a sua importância (CONSELHO FEDERAL DE PSICOLOGIA, 2010).

Também é do ano de 2003 a Resolução CFP 002/2003 que regulamentava o uso, elaboração e comercialização de testes psicológicos. Foi a segunda resolução (substituiu a pioneira Resolução CFP 25/2001) a apontar os requisitos mínimos e obrigatórios para confecção e utilização dos instrumentos psicológicos, a saber: fundamentação teórica (com ênfase na definição do construto); apresentação de evidências empíricas de validade e precisão; dados empíricos sobre as propriedades psicométricas dos itens; demonstração de sistema de correção e interpretação dos escores, características da amostra, explicando o embasamento teórico e o porquê da escolha do procedimento de interpretação utilizado; clareza na apresentação dos procedimentos de aplicação e correção, a fim de garantir uniformidade e compilação das informações supracitadas, em um manual (CONSELHO FEDERAL DE PSICOLOGIA, 2003).

Em 2004, na primeira análise dos instrumentos comercializados no Brasil após o início da vigência da Resolução, dos 106 testes cadastrados no SATEPSI, 51 (48,1%) encontravam-se sem condições de uso. Já em uma segunda análise, em 2010, dos 214 instrumentos, 77 (35,9%) estavam com parecer desfavorável, 114 (53,2%)



eram favoráveis e 23 (10, 7%) estavam em processo de análise, o que demonstrou que o número de testes e a qualidade aumentaram, levando ao aquecimento no mercado, aumento nas publicações e ainda indicando um avanço no domínio das metodologias psicométricas (VIEIRA; CAMPOS, 2011; PRIMI, 2010).

Muito recentemente, esta resolução foi revogada pela Resolução CFP 009/2018, de 24 abril de 2018 (CONSELHO FEDERAL DE PSICOLOGIA, 2018). As principais mudanças neste novo documento estão claras desde o preâmbulo. Fica explícito que o principal objetivo é o estabelecimento de diretrizes para a realização de Avaliação Psicológica, enquanto a anterior era mais reducionista, já que se referia, literalmente, à testagem. Outra característica importante é o destaque dado ao SATEPSI, considerando que esse é um sistema informatizado que organiza todas as questões referentes ao que está sendo estudado (CONSELHO FEDERAL DE PSICOLOGIA, 2018). Outrossim, destaca quais as fontes de informações fundamentais e quais as complementares na realização da Avaliação Psicológica. Mais uma mudança significativa é referente ao prazo dos estudos para validade e precisão, que agora, em semelhança aos de padronização, passam a ser de 15 anos. Acerca da comercialização dos testes, essa resolução preconiza o controle que as editoras deverão manter no que diz respeito ao nome do psicólogo e números de inscrição no Conselho Regional de Psicologia e número de série do (s) instrumento (s).

Também fica claro o quanto seu conteúdo é baseado nas definições contemporâneas de validade. Percebe-se isso especialmente pelos documentos utilizados como referência, a exemplo, o mais importante texto para a Psicometria, os *Standards for Educational and Psychological Testing*. No entanto, acerca dos critérios mínimos para verificação da qualidade dos testes, embora a atual resolução apresente algumas modificações, como fornecer uma classificação melhor para instrumentos com mais estudos de qualidade, para serem aprovados os critérios ainda são de fato mínimos (CONSELHO FEDERAL DE PSICOLOGIA, 2018; SOUZA; ALEXANDRE; GUIRARDELLO, 2017). A seguir serão apresentados os conceitos básicos das principais propriedades psicométricas utilizadas para indicar a qualidade de um teste psicológico.

## 2.2 VALIDADE

O princípio por trás do teste é a tentativa de conversão das características psicológicas não observáveis, chamadas de traço latente, em comportamentos observáveis, o que é feito por meio de tarefas propostas ao avaliando, como por exemplo, responder a questionamentos, executar comandos, desenhar, contar histórias, dentre outras (PRIMI; MUNIZ; NUNES, 2009). Em posse dessas informações e em consonância com as orientações para interpretação contidas no manual, o profissional pode testar hipóteses, fazer inferências e/ou prever comportamentos em determinados contextos, mas esses dados só serão legítimos se o instrumento possuir evidências de validade (PRIMI; MUNIZ; NUNES, 2009).

A definição de validade como “a extensão em que um conjunto de variáveis realmente representa o construto a ser medido” (SOUZA; ALEXANDRE; GUIRARDELLO, 2017) esteve, classicamente, atrelada a uma condição tripartite, a saber, de acordo com Primi, Muniz e Nunes (2009):

( a ) validade de conteúdo – investigação acerca de se os itens representam todo o domínio do comportamento a ser medido;

( b ) validade de critério – a capacidade de predição de variáveis externas e critérios, que o instrumento possui; e

( c ) validade de construto – que diz respeito ao quanto as evidências sustentam os significados da interpretação dos escores.

Nessas circunstâncias, tinha-se a ideia de que o teste media o que se propõe a medir, e as interpretações eram feitas, independentemente do contexto. Assim, de acordo com Primi, Muniz e Nunes (2009, p. 244), a validade nos *Standards for Educacional and Psychological Testing*, de 1986, passou a ser “o grau em que as evidências embasam inferências feitas a partir dos escores dos testes”.

Apesar de a definição clássica ser utilizada na atualidade, vários estudos, especialmente os aportes de Samuel Messick, apontaram para uma concepção mais abrangente, principalmente por questionar que a validade de construto estava presente também nos outros dois tipos do modelo tripartite. Urbina (2007) destacou, a partir das ideias e esquemas propostos por Embretson em 1983, a noção da “validação de constructo como uma forma unitária e abrangente de expressar a abordagem científica da integração de qualquer evidência relacionada com o sentido ou interpretação dos escores de teste”. Com isso, nos *Standards* de 1999, validade passou a se referir ao

“grau em que evidência e teoria sustentam as interpretações dos escores dos testes, vinculados aos usos propostos dos testes” (PRIMI; MUNIZ; NUNES, 2009, p. 244-246).

Percebe-se nesta definição a importância que se dá para a interpretação discriminada dos escores e com evidências para cada contexto, outrossim, teoria e prática devem caminhar juntas, levando em consideração a relevância e as consequências da testagem. Validade é, portanto, na concepção contemporânea, cumulativa, ou seja, não é necessário que se tenham todas as fontes de evidências, porém quanto mais estiverem presentes, para contextos e propósitos distintos, mais qualidade possui o instrumento. Estão listadas a seguir as fontes contemporâneas de evidências de validade (PRIMI; MUNIZ; NUNES, 2009), baseadas no (a):

Conteúdo: levanta informações sobre o quanto os itens são amostras representativas e abrangentes do domínio do construto;

Estrutura interna: verifica se as correlações entre os itens, componentes ou subtestes são coerentes com a teoria. Em outras palavras, pretende-se analisar o que sustenta um teste. Segundo Ambiel e Carvalho (2017), uma das formas mais comuns de avaliar a estrutura interna é a análise fatorial que pode ser tanto exploratória, quanto confirmatória, na primeira o objetivo é a redução de uma quantidade considerável de variáveis observáveis (itens) em partes menores (fatores), isso por meio de correlações que desembocam em agrupamentos, sendo interpretadas pela teoria previamente analisada. Já o segundo tipo mencionado é realizado quando já se tem clareza da relação esperada entre as variáveis, assim, como o nome sugere, está ligada ao teste hipóteses. Além dessas, há outras diversas formas de avaliar a estrutura interna como a Análise de Componentes Principais e Análises de Rasch.

Variáveis externas: diferente das correlações na estrutura interna, nesta categoria, levantam-se informações acerca das correlações entre os escores do teste e outras variáveis. Esse tipo de fonte de evidências pode ser de *critério* (quando se relaciona o teste a uma variável externa não psicológica) e de *convergência ou discriminante* (quando a variável externa, comparada com o teste, é também um construto psicológico). A validade de critério ainda se subdivide em *preditivo* (quando os escores podem predizer um comportamento no futuro) e *concorrente* (quando a comparação é concomitante, em geral com alguma variável que já ocorreu); a convergência diz respeito a correlações fortes, quando se utiliza testes com o mesmo construto e correlações moderadas, quando os construtos são relacionados. Já quando se

correlacionam estatisticamente construtos não relacionados teoricamente – caso da verificação da validade divergente/discriminante -, espera-se uma que a correlação seja fraca.

Processo de resposta: investiga os processos mentais em relação ao modo como se responde, ou seja, se existem variáveis que interferem no responder às tarefas do teste;

Consequências da testagem: examina quais os benefícios intencionais e não intencionais para a sociedade, decorrentes da utilização do instrumento. Uma forma de investigar essa fonte de validade seria realizar estudos experimentais, coletando dados acerca das influências de intervenção por meio dos escores obtidos no instrumento antes e depois dos procedimentos.

Tendo em vista as razões aqui explicitadas, bem como a qualidade do levantamento de informações, no período em que vigorou a Resolução 002/2003 (2003-2018), era indispensável a constância de estudos sobre os instrumentos desenvolvidos, pois, no artigo 14 do documento era indicado que “os dados empíricos das propriedades de um teste psicológico devem ser revisados periodicamente, não podendo o intervalo entre um estudo e outro ultrapassar 20 (anos), para os dados referentes a validade e precisão” (CONSELHO FEDERAL DE PSICOLOGIA, 2003, p.6). Se não fossem obedecidos os critérios estabelecidos, o teste não poderia ser legalmente utilizado na prática profissional dos psicólogos (CONSELHO FEDERAL DE PSICOLOGIA, 2003).

### 2.3 PRECISÃO

Além de saber se o teste possui evidências de validade, verificar a estabilidade desse instrumento também diz respeito à sua qualidade. Ou seja, averiguar se o instrumento é capaz de “reproduzir um resultado de forma consistente no tempo e no espaço, ou a partir de observadores diferentes” (SOUZA; ALEXANDRE; GUIRARDELLO, 2017, p. 650). Em outras palavras, a precisão refere-se ao grau de similaridade ou flutuação dos escores em momentos distintos. (SOUZA; ALEXANDRE; GUIRARDELLO, 2017).

A precisão trata da consistência da medida, somente, sem se preocupar exatamente com o que se está medindo (HOGAN, 2006), o que seria papel da validade. O critério em ênfase, também é conhecido por confiabilidade, fidedignidade, consistência interna ou estabilidade (ZANON; FILHO, 2015; SOUZA; ALEXANDRE;

GUIRARDELLO, 2017). De acordo com Zanon e Filho (2015), avaliar evidências de validade é dispendioso e demanda por procedimentos complexos, assim, a fim de diminuir os custos, alguns pesquisadores recorrem em primeira análise, à precisão, pois se não houver evidências desse critério, tampouco haverá daquele.

Outrossim, a precisão é importante para distinguir as graduações de características (habilidades, personalidade, dentre outras) dos testandos. Ademais, replicar estudos é um dos objetivos principais da ciência e se esse instrumento não possui uma mensuração adequada e estável, os dados não servirão ao fim almejado (ZANON; FILHO, 2015). No entanto, é importante deixar claro que a precisão não é suficiente para o uso de um teste, ambos os aspectos (e os demais, posteriormente explicitados) devem caminhar juntos, conforme previsto nos *Standards Educational and Psychological Testing* (1999), como era previsto pela Resolução 002/2003 do CFP e segue na Resolução 009/2018, do mesmo órgão.

Hogan (2006) ao convidar à reflexão acerca das mudanças reais no traço e as flutuações nos escores relacionadas a circunstâncias pessoais ou passageiras levanta a questão de que toda mensuração está sujeita a erros (ZANON; FILHO, 2015). Levando isso sempre em consideração, o trabalho da precisão é demonstrar que a quantidade de erros é a menor possível. O escore do erro (EE) pode ser expresso na diferença entre valor bruto obtido [escore observado (EO)] e valor que deveria ser recebido pelo testando, mas que é desconhecido [escore real (ER)] (ZANON; FILHO, 2015). A fórmula do erro seria, portanto:

$$EE = EO - ER.$$

Para Pasquali (2010), os erros podem ser próprios da observação, da amostragem, ou dos eventos em que se realizou a medida, e considera quatro fontes principais de erros de observação:

- ( a ) Instrumentais: decorrentes do instrumento utilizado;
- ( b ) Pessoais: ocorrentes pelas diferenças de reações individuais;
- ( c ) Sistemáticos: erros que não podem ser controlados, provém a algum fator sistemático;
- ( d ) Aleatórios: erros que não possuem a causa conhecida.

Há ainda os erros de amostragem, causados na seleção da amostra e que podem ocasionar inferências errôneas e falta de representatividade sobre toda a população de eventos ou objetos (PASQUALI, 2010).

A seguir são apresentados os tipos mais comuns de precisão:

Estabilidade temporal: remete à estimativa da consistência das repetições nas mensurações. O método utilizado para a verificação é o teste-reteste, que consiste na reaplicação de um instrumento, no mesmo indivíduo, em outro momento. O critério satisfatório para a interpretação dos resultados é uma correlação de, no mínimo, 0,70 entre as medidas dos dois momentos temporais (SOUZA; ALEXANDRE; GUIRARDELLO, 2017). Dentre os problemas deste tipo destaca-se o aprendizado e as mudanças reais.

Formas alternadas/ paralelas: assemelha-se ao teste-reteste, todavia, as reaplicações ocorrem com conjuntos diferentes de itens. Para isso é necessário que haja formas equivalentes do mesmo teste, ou seja, com quantidades iguais de itens, mesmo formato, dificuldade, atratividade, instruções e domínios, a exemplo: mesma quantidade de alternativas falsas e mesmo número de pontos na escala (ZANON; FILHO, 2015). Para esta análise, utiliza-se correlação, em geral de Pearson, dos escores obtidos nos dois testes, que podem ser aplicados sucessivamente um ao outro (HOGAN, 2006). Problemas em se utilizar esse método podem ser o cansaço dos participantes por terem que responder dois instrumentos e a dificuldade de ter dois instrumentos que sejam realmente equivalentes – o que pode ser solucionado com a separação do mesmo teste em duas partes, desde que haja elevada correlação entre elas e a verificação se o teste pode ser dividido sem que haja predominância de aspectos em uma parte em detrimento da outra (ZANON, FILHO, 2015).

Consistência interna/ homogeneidade: averigua se todas as subpartes de um teste estão medindo a mesma característica. O método mais comum de medida é o alfa de Cronbach (SOUZA; ALEXANDRE; GUIRARDELLO, 2017). Os valores de alfa variam entre 0 a 1, quanto mais próximo de zero, menor a fidedignidade;  $0,79 > \alpha > 0,70$  é a faixa aceitável,  $0,69 > \alpha > 0,60$ , questionável,  $0,59 > \alpha > 0,50$ , são ruins, valores abaixo de 0,50 são inaceitáveis e para que um valor seja considerado excelente, deve ser superior a 0,90 (ZANON; FILHO, 2015). Em escalas com variáveis dicotômicas é mais adequado utilizar o teste de Kuder- Richardson, sendo os valores considerados ideais também aqueles próximos de um. (SOUZA; ALEXANDRE; GUIRARDELLO, 2017; CRONBACH, 1951 apud ZANON; FILHO, 2015).

Equivalência: concerne à concordância entre avaliadores (dois ou mais) quanto aos escores e a forma mais comum é a confiabilidade interobservadores. (SOUZA; ALEXANDRE; GUIRARDELLO, 2017). De acordo com Hogan (2006, p. 102) “é importante que os dois (ou mais) avaliadores trabalhem independentemente”, já

que, se influenciados, podem fornecer notas iguais ou parecidas, “o que inflacionaria o coeficiente de fidedignidade” (SOUZA; ALEXANDRE; GUIRARDELLO, 2017, p. 652). O método consiste na dupla correção (por duas pessoas, uma pessoa e uma máquina, ou duas máquinas) de um teste feito. Os critérios para interpretação mais utilizados são o coeficiente Kappa, que assume o valor máximo de 1,00 (SOUZA; ALEXANDRE; GUIRARDELLO, 2017), que segundo Borsa e Seize (2017, p. 23) é “a razão da proporção de vezes que os especialistas concordam com a proporção máxima de vezes que deveriam concordar”, é utilizado apenas com dados categóricos e em escala nominal; o Coeficiente de Correlação Intraclasse (CCI), utiliza-se para medir a homogeneidade de duas ou mais medidas, usado na área da saúde, para verificar a concordância entre especialistas (BORSA; SEIZE, 2017).

## 2.4 NORMAS

Os valores brutos obtidos em uma testagem têm pouco propósito se não estiverem diante de uma comparação. O objetivo de estabelecer normas é atribuir significados a estes resultados, ou seja, instituir estratégias para interpretar os escores de maneira coerente e com referencial (PACICO, 2015).

Esse referencial é estabelecido a partir do desempenho da grupo/amostra normativo (a) - um grupo representativo da população a qual se destina a pesquisa, que seja aleatório e com distribuição normal, e que essa parcela da população seja representativa, a fim de que os resultados sejam razoáveis. (PACICO, 2015; URBINA, 2007). A partir daí, todos os escores individuais obtidos poderão ser comparados a esse grupo, firmando um padrão, o que ocorre a partir da conversão dos escores obtidos em percentis ou valores que utilizam distâncias entre o avaliando e o grupo de referência. Com isso, além de outros aspectos, garantem-se princípios éticos às pessoas submetidas ao método da testagem.

A lógica subjacente a esse sistema de interpretação é que os construtos psicológicos se distribuem de formas variadas (e grande parte segue a distribuição normal) e que a comparação do resultado obtido por um indivíduo com o de uma amostra dessa população permite algumas inferências sobre características mais ou menos destacadas de seu funcionamento psicológico (CONSELHO FEDERAL DE PSICOLOGIA, 2010, p. 123).

As principais normas utilizadas são:

Percentis: muito utilizado na interpretação dos testes psicológicos, diz respeito à posição do escore bruto (obtido pelo participante em comparação com o escore da amostra normativa) e indica a porcentagem de pessoas do grupo normativo que se encontra abaixo da pessoa que realizou o teste (PACICO, 2015). É importante ressaltar que o aumento no valor, não indica necessariamente o aumento do percentil. (MIGUEL, 2017). Exemplo: se o participante observado obtém escore que equivale a um valor percentílico de 30, quer dizer que 30% das pessoas do grupo de referência apresentam desempenho inferior que o seu.

Escore-padronizados: são fontes de normas em que os escores passam por distribuição linear, e são transformados em escalas com a posição em relação a uma média. (PACICO, 2015). Subdividem-se principalmente em:

a) Escore Z: concernente à interpretação a partir do desvio-padrão na curva de distribuição normal, a média é equivalente a 0 e o desvio padrão igual a 1. A distribuição é simétrica e bilateral, ou seja, o valor do sinal (negativo ou positivo indica o sentido de deslocamento em relação à média. É calculado da seguinte forma:

$$Z = \frac{\text{escore bruto} - \text{média da amostra normativa}}{\text{Desvio padrão da amostra normativa}}$$

(PACICO, 2015, p. 50)

b) Escore T: a fim de evitar o sinal negativo, algumas transformações podem ser feitas no escore-padrão, um exemplo é o escore T, em que se multiplica o valor do escore (padronizado) a um número e soma-se uma constante ao resultado:

$$T = 50 + 10 Z. \text{ (PACICO, 2015, p. 50 e 51)}$$

( C ) Quociente de inteligência: Seguindo a mesma lógica do escore T, está o QI atribui-se 100 à média, e 16 (ou 15, de acordo com Miguel, 2017, p. 130) ao desvio padrão:

$$QI = 100 + 16 Z. \text{ (PACICO, 2015, p. 51)}$$

Para Pacico (2015), além das fontes de normas baseadas na amostra normativa, há a produção de normas baseadas no desenvolvimento humano, que são as:



normas de desenvolvimento, normas de idade, normas de série escolar, normas de estágio de desenvolvimento, referentes ao estágio psicomotor em que a criança se encontra. Testes foram desenvolvidos com base nas fases sugeridas por Piaget para o desenvolvimento: sensório motor, pré-operacional, operacional concreto e operacional formal, essas seguem o mesmo princípio comparativo, entre o valor encontrado e a comparação com os resultados típicos do grupo para cada característica.

Segundo Miguel (2017), além das interpretações baseadas em expectativas normativas e/ou distribuição, existem aquelas baseadas em critério e domínio, a primeira diz respeito à localização do escore em uma nota de corte ou base, por exemplo, em uma escala; já o segundo tipo é mais encontrado em testes de desempenho cognitivo, baseados na maestria do indivíduo, no modo como este alcança ou não o que é esperado, esse tipo de instrumento é mais encontrado em ambientes escolares e universitários. Nos testes de autoexpressão ou projetivos, o que mais se encontra em relação às normas são categorias que configuram o perfil do testando.

Assim como para validade e precisão, a Resolução CFP 002/2003 (CONSELHO FEDERAL DE PSICOLOGIA, 2003, p. 6) considerava que “os dados empíricos devem ser revisados periodicamente, não podendo o intervalo entre um estudo e outro ultrapassar 15 (quinze) anos, para os dados referentes à padronização”, que inclui as normas. Na nova Resolução, não houve alteração nesse período de tempo.

### 3 MÉTODO

O método de investigação deste estudo tem como ponto de partida a publicação da Resolução CFP 002/2003 e, de chegada, o ano de 2017, imediatamente anterior ao de início da vigência da Resolução CFP 009/2018, a fim de que pudesse ser realizado o comparativo entre os anos das publicações das duas resoluções. O método utilizado foi a Revisão Integrativa (RI) que, segundo Mendes, Silveira e Galvão (2008), consiste em estabelecer objetivos da pesquisa e gerar hipóteses, realizar a busca de materiais, levando em consideração os critérios de inclusão e exclusão, com a finalidade de reunir e sintetizar os dados para aprofundar o conhecimento e, a partir daí, formular conclusões gerais, tanto para contribuir com discussões quanto para identificar as lacunas que precisam ser investigadas na área.

#### 3.1 BUSCA NA LITERATURA E CRITÉRIOS DE SELEÇÃO DOS ARTIGOS

Foram consultados seis especialistas em Psicometria sobre quais seriam as palavras-chave mais adequadas para realizar a busca dos artigos que pudessem compor a amostra, de acordo com os objetivos propostos para o estudo. A partir dos resultados dessa etapa, as seis palavras-chave mais indicadas foram aplicadas na base de dados *Scielo*. Nesta base, primeiramente foi feita uma busca sem filtros, apenas para quantificação dos resultados gerados. Em seguida, foram aplicados filtros, em que as publicações selecionadas obedeceram aos critérios de inclusão: (a) serem referentes aos últimos 15 anos (2003 a 2017), período de vigência da Resolução CFP 002/2003; (b) apresentassem instrumentos psicológicos desenvolvidos para o contexto brasileiro; (c) escritas em língua portuguesa ou inglesa, uma vez que é relativamente comum a publicação de artigos nacionais nessa língua; (d) pertencentes à área temática WOS (*Web of Science*) - Psicologia e (e) pertencentes ao tipo de literatura “artigo”. A cada nova palavra-chave utilizada, os títulos foram analisados para verificar se não havia duplicidade de artigos. Em seguida os resumos foram lidos a fim de extrair a frequência com que cada propriedade psicométrica era utilizada. Portanto, as palavras-chave foram apenas o meio para investigar a frequência das três propriedades psicométricas propostas no objetivo: validade, precisão e normas.

### 3. 2 APRESENTAÇÃO DOS RESULTADOS

Além da análise geral da frequência de utilização de cada propriedade psicométrica, foi feita uma análise comparativa conforme três períodos de tempo, estabelecidos para esta pesquisa: (a) 2003 -2007 – fase de implementação da Resolução CFP 002/2003; (b) 2008 – 2012 – fase de adaptação da referida Resolução; (c) 2013 – 2017 – fase de consolidação, imediatamente anterior à Resolução CFP 009/2018.

#### 4 RESULTADOS E DISCUSSÃO

Nesta seção são apresentadas e comentadas as propriedades psicométricas encontradas nos resumos dos artigos selecionados para este trabalho. Primeiramente são verificadas quais as palavras-chave mais indicadas pelos juízes. Em seguida, são mostrados os dados referentes à validade, com comentários acerca das fontes preconizadas pelos *Standards for Educacional and Psychological Testing*, como também das exceções identificadas, que fogem desta orientação. Depois são discutidos os achados acerca da precisão e normas, nos mesmos moldes do critério anterior e, por fim, apresentadas informações acerca da frequência de produção por regiões do Brasil.

Os resultados em relação às palavras-chave indicadas pelos juízes como mais adequadas para alcançar os objetivos do estudo estão apresentados na Tabela 1, a seguir.

**TABELA 1 – Palavras-chave sugeridas pelos juízes especialistas**

<b>J</b>	<b>A</b>	<b>T</b>	<b>PALAVRA 1</b>	<b>PALAVRA 2</b>	<b>PALAVRA 3</b>	<b>PALAVRA 4</b>	<b>PALAVRA 5</b>
1	2008	7	Psicometria	Avaliação Psicológica	Validade	Normatização	Confiabilidade/ Precisão
2	2010	8	Psicometria	Validação	Precisão	Fidedignida- de	Medida
3	2006	8	Content validity	Structural validity	Internal consistency	Norms	Semantic analysis
4	2006	14	Validade	Fidedignidade	Padrões normativos	Validade de construto	Confiabilidade
5	2009	11	Propriedades psicométricas	Validade e fidedignidade	Validação da versão brasileira	Adaptação brasileira	Teste brasileiro
6	2009	8	Construção de instrumentos	Adaptação de instrumentos	Validade	Consistência interna	Normatização de instrumentos

Elaboração da autora, com base nos dados da pesquisa.

NOTA: J – Juiz; AF - Ano de formação; T – Tempo de experiência em Psicometria

Em linhas gerais, percebe-se que não há um consenso sobre quais termos seriam os mais corretos. Com o agrupamento, conforme apresentado na Tabela 2, foi possível observar que termos relacionados à validade apareceram seis vezes; fidedignidade, precisão e confiabilidade, juntas, também totalizaram seis; normas apareceram quatro vezes e psicometria foi indicada por dois especialistas. No entanto, há diferentes formas de nomear as propriedades psicométricas dos instrumentos.

**TABELA 2 - Frequência das palavras- chave por semelhança**

Psicometria	2
Validade	2
Validação da versão brasileira	1
Validade de construto	1
Content validity	1
Structural validity	1
Fidedignidade	2
Confiabilidade	2
Precisão	2
Normatização	1
Normatização de instrumentos	1
Padrões normativos	1
Norms	1
Avaliação psicológica	1
Construção de instrumentos	1
Adaptação de instrumentos	1
Consistência interna	1
Internal consistency	1
Adaptação brasileira	1
Propriedades psicométricas	1
Teste brasileiro	1
Semantic analysis	1

Elaboração da autora, com base nos dados da pesquisa.

Os resultados seguintes indicaram mais de 15 mil resultados brutos que, com os filtros, foram reduzidos a 362 artigos (ver nos apêndices a relação completa dos artigos), com maior concentração na palavra-chave validade, seguida por psicometria, fidedignidade, precisão, normas e confiabilidade, nessa ordem. Informações detalhadas podem ser observadas na Tabela 3, a seguir.

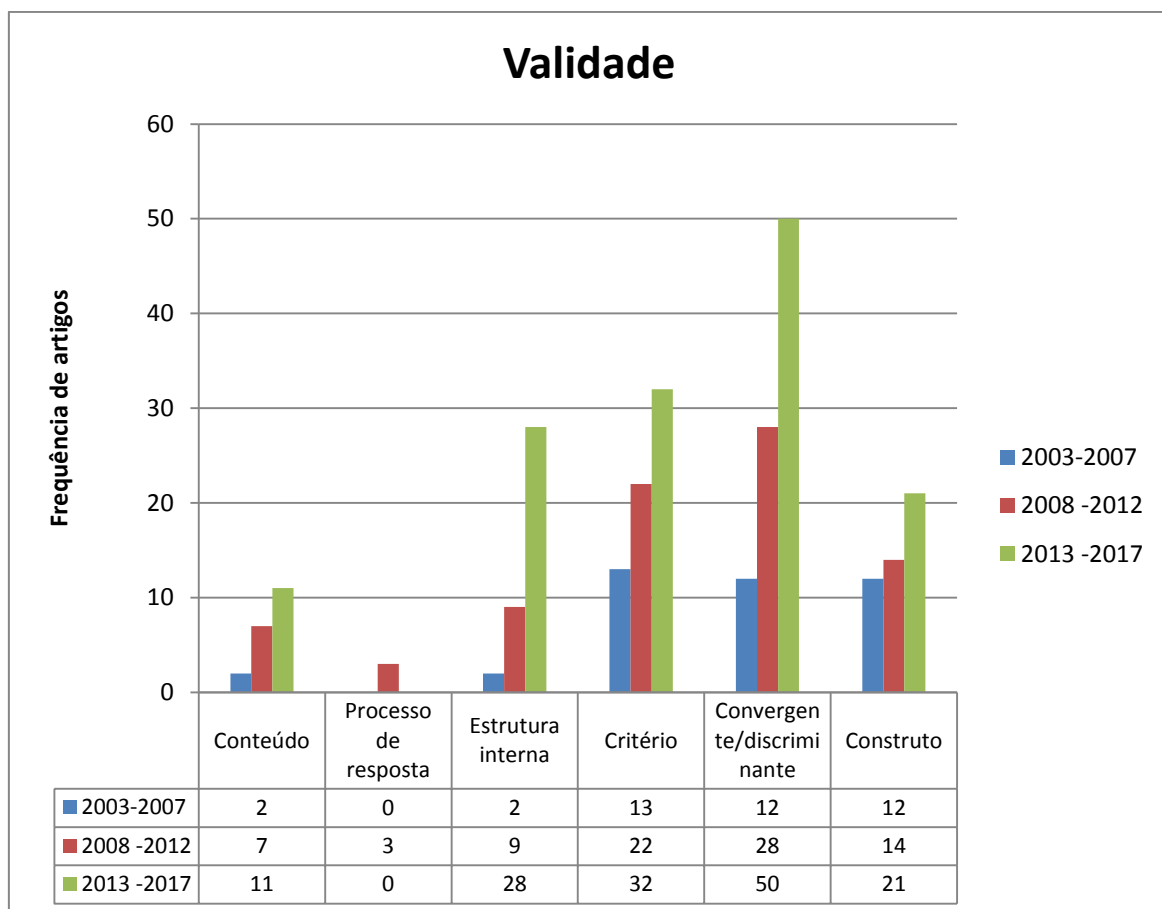
**TABELA 3 - Quantidade de artigos por categorias**

<b>PC</b>	<b>OB</b>	<b>SF</b>	<b>CF</b>	<b>NE</b>	<b>RP</b>	<b>RS</b>
Psicometria	1°	608	92	22	-	69
Validade	2°	3415	373	82	47	244
Fidedignidade	3°	321	63	14	31	18
Precisão	4°	2952	102	34	54	14
Normas	5°	5893	107	87	9	11
Confiabilidade	6°	2432	24	6	10	6
<b>Total</b>		<b>15.621</b>	<b>761</b>	<b>245</b>	<b>151</b>	<b>362</b>

Elaboração da autora, com base nos dados da pesquisa.

Nota: PC – palavra – chave; OB – Ordem de Busca; SF – Quantidade sem filtros; CF – Quantidade com filtros; NE – Artigos que não se encaixaram na pesquisa; RP – Repetidos; RS - Restantes

As frequências totais encontradas para as fontes de validade foram: 3 para validade de processo de resposta, 20 de conteúdo, 39 de estrutura interna, 47 para validade de construto, 67 de critério (concorrente e preditivo) e 90 de evidências de validade convergente ou discriminante. O Gráfico 1, apresentado a seguir, é um comparativo dos totais entre as principais fontes de validade encontradas na pesquisa, divididas pelos três períodos selecionados. Como apresentado anteriormente, a definição contemporânea menciona fontes de validade, que acumuladas fornecem o grau de evidências de validade de um instrumento (PRIMI; MUNIZ; NUNES, 2009). Portanto, mais que uma possibilidade, é desejável que um mesmo instrumento tenha evidências de diferentes fontes de validade.

**GRÁFICO 1 - Comparativo entre as principais fontes de validade por período**

Elaboração da autora, com base nos dados da pesquisa.

Observando o gráfico, percebe-se que, exceto para Validade de Processo de Resposta, houve um aumento por período no número de estudos em relação à validade dos instrumentos no Brasil. Primi e Nunes (2010) afirmaram que a Resolução CFP 002/2003 surgiu como uma resposta ao número elevado de processos éticos envolvendo a Avaliação Psicológica e, com essa implementação, uma série de mudanças ocorreram e isso induziu a melhoria na qualidade dos testes. Além disso, a expectativa dos autores naquela época era de avanços mais promissores nesse aspecto. Bueno, Amorim e Albuquerque (2017) relatam algo semelhante. Segundo eles, a implantação do SATEPSI, a Resolução CFP 002/2003 e a consequente regulamentação na área trouxeram efeitos benéficos, tendo em vista a maior produção de pesquisas na área. Desse modo, os dados da presente pesquisa reforçam essa tendência prevista.

Percebe-se que os maiores valores de frequência, no último período e no total, dizem respeito aos estudos da chamada validade baseada na relação com variáveis externas, que contemplam tanto validade de critério (concorrente e preditivo), quanto



validade convergente (por mesmo construto ou construtos relacionados) e validade discriminante ou divergente, em que nas três divisões de tempo, apresentam quantidades significativas. Ao todo foram 67 artigos que mencionaram a fonte critério e 90 que se referiram à convergência/ divergência. De acordo com Freitas e Damásio (2017) as evidências de validade que tem por base a relação com variáveis externas, devem ser investigadas com frequência, já que estão intrinsecamente ligadas aos contextos nos quais foram encontradas. Assim, são fundamentais para mostrar que um teste pode ser usado com maior segurança em diferentes contextos e com diferentes propósitos, conforme sustenta a definição contemporânea de validade.

Outro aspecto que chama atenção nesses resultados é o uso dos termos “validade de construto” e “validade de estrutura interna”. Nomear validade de construto de maneira independente é um pressuposto que remete à definição clássica de validade (PRIMI; MUNIZ; NUNES, 2009). A partir dos anos 1980, com as discussões propostas principalmente por Samuel Messick, ganhou força o que nos anos 1990 viria a ser chamada de definição contemporânea de validade (PRIMI; MUNIZ; NUNES, 2009). A partir desta definição, passou-se a entender que todas as investigações sobre validade são em relação ao construto, não devendo haver uma específica e independente com esse nome. Por isso, nessa nova concepção, a indicação passou a ser o uso do termo “validade de estrutura interna”.

Assim, a presença de 47 artigos ao longo desses 15 anos utilizando o termo validade construto, sendo 21 no período atual (2013 – 2017), conduz à reflexão, pois demonstra que a presença da definição clássica persiste. No entanto, nota-se também que a validade de estrutura interna cresceu em grande proporção ao longo dos anos, o que já pode ser uma consequência da utilização das definições contemporâneas de validade pelo Conselho Federal de Psicologia em seus documentos e resoluções.

Ainda sobre estrutura interna, os artigos trouxeram aspectos relacionados à utilização da técnica da análise fatorial. No período compreendido entre 2003 e 2007, foram feitas 44 menções à Análise Fatorial, já nos anos de 2008 a 2012 ela apareceu em 53 dos resumos, e, por fim, de 2013 a 2017, 98 das pesquisas a usaram, demonstrando um crescimento ao longo dos períodos. A análise fatorial, segundo Ambiel e Carvalho (2017) é uma das formas mais comumente utilizadas na verificação de estrutura interna.

Mais um ponto sobre o tema que pode ser destacado é a maior quantidade de estudos que indicam nominalmente a estrutura interna em detrimento aos que investigaram conteúdo. Considerando que esta fonte de validade, pela lógica, deveria

ser observada antes da estrutura interna, como propõe Pasquali (2010), é surpreendente que poucos estudos se preocupem com evidências de validade de conteúdo. Isso pode ocorrer em consequência da maior valorização dada à verificação da coesão entre os itens do que à representatividade e abrangência dos mesmos. A própria Resolução CFP 009/ 2008, no item B 7.2 preconiza que é “indispensável a apresentação de um estudo relacionado à validação da estrutura interna do teste”. Considerando que analisar o conteúdo do teste antes da estrutura poderia ajudar a fornecer estruturas mais coesas e estáveis, de certa forma, os resultados indicam que os pesquisadores parecem assumir riscos ao abrir mão deste procedimento.

Por fim, chama a atenção a ausência de estudos sobre consequências da testagem e a baixa frequência no que se refere às evidências de validade de processo de resposta. Apenas 3 artigos mencionam a utilização desta fonte e todos estão no período compreendido entre os anos de 2008 a 2012. De acordo com Primi, Muniz e Nunes (2009) essas fontes são as mais recentes. Processo de resposta aparece a partir da década de 1980, em decorrência da Psicologia Cognitiva e é mais frequentemente utilizada em testes de inteligência. Sobre as consequências da testagem, no meio científico há discordâncias acerca dessa fonte, já que não diz respeito à estrutura do teste em si, mas a um caráter ético que perpassa o aspecto da vinculação das consequências que foi e está sendo gerada com a utilização dos instrumentos. Os resultados indicam, portanto, que há dois campos ainda pouco explorados das evidências contemporâneas de validade. Dessa maneira, embora teoricamente ela seja formada por cinco fontes de validade, na prática, uma estrutura tripartite continua vigente, embora com alguns refinamentos em relação à divisão clássica.

Além destas fontes de evidência de validade destacadas por sua frequência, outros termos aparecem nos artigos, como validade de face, validade aparente, validade nomológica e validade de grupos comparados. Esses dados podem ser observados nas cinco tabelas apresentadas nos Apêndices, referentes às fontes de validade encontradas dentro dos resumos dos artigos, ao buscar cada uma das seis palavras-chave mais indicadas pelos juízes para a pesquisa, divididas por período. Tais tabelas são importantes também porque contém os dados do Gráfico 1 mais detalhados.

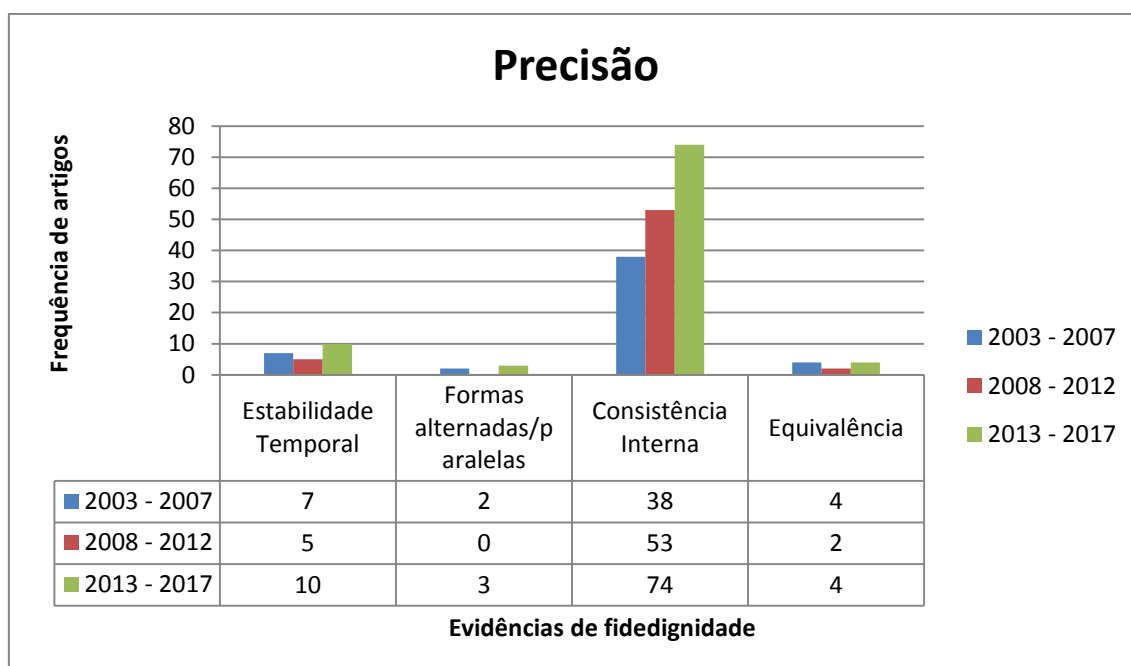
Sobre esse aspecto dos diferentes nomes dados para as validades, Pasquali (2007) tece a crítica de que a criação e existência de diferentes definições e tipos de validade gera confusão. O autor afirma que a validade constitui uma propriedade do instrumento e não tem a ver com o julgamento que se faz dos seus escores, assim o teste

é válido ao medir o que supostamente deve medir. Considera que validade vai responder se algo é verdadeiro ou falso e precisão é que vai dizer se algo está errado ou correto, ou seja, diz que não tem sentido considerar que um instrumento é válido em um contexto, mas não o é no outro, isso é um problema de “calibragem” do instrumento, mais relacionado à precisão e não à validade. Com isso, afirma que tanto validade quanto precisão são importantes, mas que um não é o outro. E que com a quantidade de termos, atualmente, validade é tudo e não explica mais nada.

No entanto, o próprio Conselho Federal de Psicologia, especialmente na resolução CFP 009/2018, segue os documentos que levam em consideração a definição contemporânea de validade, a saber, o próprio *Standards for Educational and Psychological Testing*, entre outros. Assim, percebe-se a importância de questionar se os instrumentos são estanques, ou seja, válidos independente dos contextos, e ter um esforço continuado para analisar as evidências de validade, já mencionadas.

Além de validade, outra propriedade investigada nessa pesquisa foi a precisão dos instrumentos. Acerca disso, não há consenso sobre qual dos termos seria mais adequado: fidedignidade, precisão ou confiabilidade. Nessa pesquisa, percebeu-se que o termo que aparece com mais frequência é precisão, com 58 usos, seguido de fidedignidade, utilizado 43 vezes e de confiabilidade, aparecendo 33. Com fins de padronização, será utilizado nesta discussão o termo mais frequente: precisão. A partir desses dados, sugere-se que este seja o termo adotado nas pesquisas e manuais de instruções, por entender que é o termo que mais sugere a característica da consistência pela diminuição de erros. Ser preciso é ser certo, correto, devidamente calibrado para o uso, o que vai ao encontro da característica mais importante do critério agora abordado, ser o mais consistente e livre de erros possível, em diferentes formas de medição.

Sobre os tipos de precisão, as frequências totais encontradas para estudos sobre os tipos de precisão ao longo dos últimos 15 anos foram: 5 para formas paralelas/alternadas, 10 para equivalência entre avaliadores, 22 para estabilidade temporal e 164 vezes com consistência interna. O Gráfico 2, apresentado a seguir, demonstra o comparativo entre os principais tipos de precisão encontrados ao longo dos três períodos adotados neste trabalho.

**GRÁFICO 2 - Comparativo entre as principais fontes de precisão**

Elaboração da autora, com base nos dados da pesquisa.

Por meio da análise do gráfico, percebe-se que a maior quantidade de estudos concentra-se na avaliação da consistência interna, com crescimento constante ao longo dos três períodos, nos últimos 15 anos. E dentro dele, o método mais comumente utilizado é o do coeficiente alfa ou alfa de Cronbach. Entre 2003 e 2007 – foram 28 estudos que mencionaram o uso de alfas de Cronbach, de 2012 a 2008, 35, e 2013 a 2017 – 37 menções a ele (os dados detalhados estão do Apêndice F ao J). O alfa de Cronbach, usado desde a década de 1950 para verificar precisão, é utilizado pela maioria dos pesquisadores para investigar a consistência interna dos instrumentos (SOUZA; ALEXANDRE; GUIRARDELLO, 2017). Como se percebe, o seu uso é tradicional, porém muitos estudiosos utilizam-no como uma única fonte para investigação da consistência interna, adotando de maneira mais tímida os outros métodos existentes.

Ainda sobre consistência interna, outro aspecto que merece destaque é que ao se investigar dados acerca dessa precisão, outra possibilidade é utilizar o método das metades, em que a correlação entre as metades deve ser obtida e então corrigida por meio do coeficiente de correlação Spearman-Brown. Entretanto, os dados obtidos neste estudo demonstram que o método não está sendo utilizado ou pelo menos seu uso não está sendo mencionado nos resumos dos artigos.

Para Urbina (2007, p. 133) “as medidas de consistência são procedimentos estatísticos que procuram avaliar a extensão da inconsistência entre os itens de um teste” e frisa que o método das metades, até determinado ponto, realiza essa tarefa. No entanto, os coeficientes alfa e Kuder-Richardson, por serem fórmulas de correlação entre todos os itens, têm maior êxito nesse âmbito, já que mesmo um teste curto tem várias possibilidades de comparação. A pouca quantidade de dados referentes à utilização do coeficiente de Kuder-Richardson pode ser um indicativo da pouca pesquisa com instrumentos dicotômicos (SOUZA; ALEXANDRE; GUIRARDELLO, 2017). Neste estudo, apenas 1 estudo entre 2017 e 2013 e 2 entre 2012 a 2008 mencionaram o uso desse coeficiente.

O segundo método mais citado para investigar a precisão de um instrumento foi a estabilidade temporal, mencionada 22 vezes. A diferença de frequência entre este método e consistência interna pode se dar porque a estabilidade temporal demanda maior investimento, em vários aspectos, como por exemplo, o tempo entre as aplicações e a disponibilidade dos testandos. A respeito do tempo, Urbina (2007) discorre que não há um intervalo fixo recomendado entre as aplicações. Se o intervalo for muito curto, os testandos podem lembrar as respostas, já se for longo experiências entre as aplicações podem influenciar nos escores, aumentando o risco de erros aleatórios. Determinar um tempo entre as aplicações, pode ser, portanto, um dificultador. Outro, relacionado a isso, está o fato de que testar estabilidade temporal implica em encontrar os mesmos sujeitos, em tempo igual ou muito semelhante, com disponibilidade para realizar a mesma tarefa já realizada anteriormente.

Por fim, com pouca frequência, aparecem estudos sobre precisão de formas alternadas ou paralelas, de equivalência entre avaliadores e outras ainda mais raras. As frequências com que esses tipos de precisão podem ser observadas nas tabelas disponibilizadas na seção dos Apêndices. A baixa frequência destes outros métodos pode ser explicada, por exemplo, pela dificuldade em encontrar instrumentos que sejam correspondentes, com quantidades iguais de itens, mesmo formato, dificuldade, atratividade, instruções e domínios, a exemplo, mesma quantidade de alternativas falsas e mesmo número de pontos na escala (ZANON; FILHO, 2015), ou ainda na equivalência, que depende tanto da padronização da aplicação quanto de um treino adequado dos avaliadores.

Uma última consideração sobre a precisão, que merece atenção, é a falta de um consenso entre os valores para sua interpretação de qualidade. A literatura

recomenda valores superiores a 0,70 (SOUZA; ALEXANDRE; GUIRARDELLO, 2017) como aceitáveis, mas a atual Resolução CFP 009/2008 manteve o valor de 0,60 como um valor suficiente (nível B) que para Zanon e Filho (2015) está em uma faixa questionável ( $0,69 > \alpha > 0,60$ ); embora o documento anteriormente citado tenha valorizado índices maiores para as classificações A e A+ (CONSELHO FEDERAL DE PSICOLOGIA, 2018). É possível, então, questionar se valores tão baixos são mesmo suficientes para mostrar que um instrumento de fato é preciso.

Diante da quantidade de dados sobre validade e comparando-os aos dados relativos à precisão, questiona-se o porquê de não existirem tantas investigações sobre a precisão quanto para a validade, inclusive tendo em vista que na prática, parece fácil investigar esta quando já se possui os dados coletados para verificar validade. Validade e precisão devem caminhar sempre juntas, mas não é isso que parece estar acontecendo, já que a preocupação com a validade parece maior. Segundo Souza, Alexandre e Guirardello (2017) a precisão deve ser sempre discutida, já que depende de vários aspectos como circunstâncias, contexto e propósito, um dos motivos pelos quais os estudos sobre esse critério devem ser sempre realizados. Não investigar precisão, mesmo já tendo os dados sobre validade, implica no desconhecimento acerca do quão o instrumento é estável, equivalente ou homogêneo, de acordo com as populações às quais se investiga.

No que tange às propriedades psicométricas relacionadas às normas, dos artigos selecionados, apenas 3 especificaram nos resumos quais os tipos de normas foram utilizados, nesses casos, uso de percentis. Este dado é relevante, tendo em vista a importância de estudos acerca da normatização. Uma das hipóteses é o desconhecimento ou desvalorização em relação aos critérios de interpretação das pontuações do instrumento. É possível que, por uma falha na formação, muitos acreditem que o uso dos escores brutos é suficiente para interpretar os resultados de um teste ou se satisfaçam com eles.

Outro ponto a ser considerado quando se trata de normas é a dificuldade em coletar dados com uma amostra representativa, considerando que isso depende da cooperação de várias pessoas, de várias localidades do país, o que, na maioria dos casos, não acontece, além das dificuldades financeiras no custeio dessas pesquisas. Quando as normas desse tipo são geradas, as implicações podem ser instrumentos com tabelas que não contemplam a população de forma adequada ou construídas com grupos de pessoas sem nenhuma representatividade na interpretação dos escores.

Portanto, pode ser que os dados normativos não estejam sendo publicados porque a amostra tem um alcance tão curto, tão regional ou local, que qualquer interpretação feita além desses limites poderia ser considerada inútil. E quando utilizadas, podem ocasionar um problema técnico e ético por parte dos profissionais que, por vezes, ao não terem dados disponíveis para interpretação, encontram como alternativa aqueles com características que se aproximem, como por exemplo, utilizar uma tabela normativa do Amazonas para um indivíduo do Maranhão.

Toda essa dificuldade tem feito que a investigação e publicação de tabelas normativas esteja, atualmente, a cargo das editoras que adquirem os direitos autorais dos testes. Com maior poder financeiro e maior alcance de rede de contatos, formada tanto por profissionais distribuidores quanto por parcerias com universidades, são elas que têm feito o trabalho exigido pelo Conselho Federal de Psicologia de ter normas representativas, de tamanho elevado e distribuída em diferentes regiões do Brasil. Se por um lado isso é uma grande contribuição para a comercialização de testes de maior qualidade, por outro lado, faz com que pequenos desenvolvedores de instrumentos fiquem impossibilitados de cumprir as recomendações, praticamente sendo obrigados a cederem os direitos para as editoras se quiserem comercializar seus instrumentos.

Por fim, uma última análise refere-se à investigação das regiões onde os dados sobre as propriedades psicométricas estão sendo gerados. A Tabela 4, a seguir, mostra a quantidade de pesquisadores vinculados a instituições em cada um dos Estados, não necessariamente à quantidade de artigos publicados em cada região.

**TABELA 4 - Produção por Regiões do Brasil (Ordem Decrescente)**

REGIÃO	QUANTIDADE
Sudeste	220
Sul	95
Nordeste	60
Centro – Oeste	58
Norte	5

Elaboração da autora, com base nos dados da pesquisa.

Segundo Gouveia (2009), a Avaliação Psicológica desenvolveu-se inicialmente no eixo Brasília – São Paulo – Rio Grande do Sul o que explica a grande quantidade de pesquisas na Região Sudeste e Sul. No entanto, de acordo com o autor, que é professor na Universidade Federal da Paraíba, não demorou muito para que nos

outros Estados também tivessem profissionais comprometidos com a sua boa prática. Ele exalta a importância não apenas da comunicação entre instituições, mas também o intercâmbio entre integrantes, ou seja, seria mais adequado, para a disseminação de conhecimentos que alguém que fez graduação em uma instituição, não prossiga a sua formação exclusivamente naquele ambiente, mas que possa, dentro das possibilidades, participar de outros ambientes acadêmicos.

O mesmo Gouveia (2009) afirma que existem pesquisas específicas nessa área – que permeia a prática de todo psicólogo – em vários estados do Nordeste, como Rio Grande do Norte, Alagoas, Ceará e Piauí, estudos coordenados por profissionais que se formaram no eixo Brasília- Sudeste – Sul e se deslocaram para a região Nordeste. A Tabela 5 traz os dados referentes à quantidade de artigos com pessoas vinculadas a instituições nos Estados desta região.

**TABELA 5 - Produção por Estados do Nordeste**

ESTADO	QUANTIDADE
Paraíba	34
Rio Grande do Norte	9
Bahia	8
Pernambuco	7
Ceará	5
Piauí	5
Alagoas	4
Sergipe	1
Maranhão	0

Elaboração da autora, com base nos dados da pesquisa.

A Paraíba lidera em número, e com uma diferença significativa em relação aos próximos, que são Rio Grande do Norte e Bahia. O Maranhão é o único estado do Nordeste que não possui nenhum artigo publicado, dentre os que foram pesquisados. Importante destacar que os dados são referentes a pessoas vinculadas a estes Estados, no entanto, em alguns casos, os artigos não são exclusivos do Nordeste e em outros, foram participantes da pesquisa pessoas de mais de um Estado, por isso a diferença entre a quantidade de artigos na região e a soma da quantidade por estados.

A Região Norte é a que apresenta o menor número de pessoas que produziram material na temática. De acordo com a pesquisa realizada por Freires et al.



(2017) os principais problemas enfrentados no Norte Brasileiro são: escassas evidências de literatura sobre processo de formação na região, diversidade de nomenclatura nas disciplinas de Avaliação Psicológica, cursos de graduação que ofertam em média apenas 3,82 disciplinas na área, sendo que essas são mais voltadas ao psicodiagnóstico e ao tecnicismo. Além disso, a pesquisa demonstra baixa carga horária destinada à Avaliação Psicológica, sendo destinadas 240 horas de uma média total de 4224 horas da carga horária total dos cursos. É, portanto, um problema que reflete não apenas na produção, mas também na prática profissional, tendo em vista que quase todo psicólogo realiza Avaliação Psicológica.

## 5 CONSIDERAÇÕES FINAIS

Os resultados deste estudo mostraram um perceptível aumento na investigação de características psicométricas dos instrumentos ao longo dos últimos 15 anos, principalmente no que diz respeito à validade, o que sugere uma melhoria nas suas qualidades. No entanto, o caminho a ser percorrido ainda é longo, tendo em vista que algumas das propriedades psicométricas não estão sendo investigadas corretamente ou ainda que alguns aspectos precisam ser refinados, como a mescla entre definições clássica e contemporânea de validade, estudos com a fonte de validade de Processo de Resposta e ausência de pesquisas baseadas na consequência da testagem, a fim de perceber quais os benefícios têm sido gerados para os indivíduos e para a sociedade.

Em face da demanda social pela qualidade da Avaliação Psicológica, são necessárias tanto melhorias profissionais, com psicólogos que atuem de maneira crítica na elaboração e proposição de mudanças, quanto na qualidade dos instrumentos psicológicos, tendo em vista que esses são fontes fundamentais no processo psicológico acima mencionado. Isso passa primeiramente pela construção de conhecimentos. O uso de instrumentos e o processo decorrente dessa utilização requerem formação básica e continuada básica e adequada.

Dentre os principais problemas referentes a isso, Borsa (2016) cita a carga horária reduzida, pouca qualificação dos professores na área e visão limitada, fragmentada e preconceituosa. Em relação às propostas para melhoria, destaca o uso de estratégias que desenvolvam habilidades, como estudos de caso, oficinas de elaboração de documentos e estágios específicos em Avaliação Psicológica. Além disso, afirma que as clínicas-escolas poderiam servir como meio para identificação de principais características e problemas prevalentes na região específica, essas informações poderiam ser fomento para Políticas Públicas.

Outro ponto a ser destacado e que tem relação direta com o que foi anteriormente considerado é a pouca disseminação da pesquisa em ambientes não acadêmicos. Considera-se que deveria haver uma maior integração tanto entre os laboratórios, grupos de pesquisa, quanto com as demais instituições e/ou editoras. Desse modo, as dificuldades em conseguir uma amostra normativa representativa seriam menores e a ciência psicológica avançaria mais rapidamente, já que dados obtidos em um local poderiam ser utilizados em outros, otimizando a pesquisa e os estudos. Com isso, seria esperado uma diminuição nos problemas de referencial para a interpretação,

ocasionando em uma substancial mudança na qualidade das evidências psicométricas e consequentemente na ética da Avaliação.

Dentre as limitações do estudo, marca-se a restrição da leitura aos resumos, em face da quantidade de artigos publicados, ao longo dos 15 anos. No entanto, frisa-se que o esperado é que os resumos contenham os dados mais significativos da pesquisa realizada. Entre os pontos fortes, está a importância de um estudo que verifique a qualidade de fontes fundamentais para a Avaliação Psicológica, a fim de se observar o que está sendo realizado em forma de investigação e melhoria dos instrumentos, após a primeira quinzena oficial da retomada das preocupações com os processos de testagem no Brasil.

Acerca dos resultados para a prática do psicólogo, espera que este estudo conduza a reflexões sobre os critérios que levam um instrumento a ser comercializado, levando em consideração que, mesmo que ele seja favorável para o uso, os critérios estabelecidos são mínimos. Que isso possa levar à leitura minuciosa do material dos instrumentos e a busca e cobrança por aperfeiçoamento. Embora o desenvolvimento caiba aos pesquisadores e a comercialização às editoras, ao profissional cabe escolher os melhores instrumentos que atendam os objetivos da sua prática ou cobrar por instrumentos de maior alcance e qualidade, de acordo com suas especificidades.

Para a agenda de pesquisa, sugere-se a identificação de artigos que utilizam validade e precisão como sinônimos; a quantificação de artigos que foram publicados, por ano; a verificação mais detalhada de critérios referentes à investigação da Análise Fatorial; a leitura do método, dos resultados e discussões, a fim de investigar aspectos mais específicos, como o tempo estabelecido nas pesquisas que usaram o teste-reteste para investigar estabilidade temporal; e a averiguação da qualidade das informações fornecidas pelos manuais dos testes. Por fim, recomenda-se fortemente que um novo levantamento seja realizado nos anos de 2023 e 2028, como forma de avaliar o impacto dos primeiros 5 anos e primeira década da nova resolução para avaliação psicológica.

## REFERÊNCIAS

- AMBIEL, R. A. M; CARVALHO, L. F. Definições e papel das evidências de validade baseadas na estrutura interna em Psicologia. In: DAMÁSIO, B. F.; BORSA, J. C (Org.). **Manual de desenvolvimento de instrumentos psicológicos**. São Paulo: Vetor, 2017. p. 85 – 100.
- AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION, NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION (AERA, APA, NCME). **Standards for Educational and Psychological Testing**. 1986.
- AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION, NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION (AERA, APA, NCME). **Standards for Educational and Psychological Testing**. 1999.
- ANASTASI, A.; URBINA, S. **Testagem Psicológica**. Porto Alegre: Artes Médicas, 2000. 575 p.
- BORSA, J. L. Considerações sobre a formação e a prática em Avaliação Psicológica no Brasil. **Temas em Psicologia**, n. 24, n 1, p. 131 – 143, 2016.
- BORSA, J. C.; SEIZE, M.M. Construção e adaptação de instrumentos psicológicos: dois caminhos possíveis. . In: DAMÁSIO, B. F.; BORSA, J. C (Org.). **Manual de desenvolvimento de instrumentos psicológicos**. São Paulo: Vetor, 2017. p. 15- 37.
- BUENO, J. M. H. A implantação do SATEPSI e seus impactos na área de Avaliação Psicológica. In: DAMÁSIO, B. F.; BORSA, J. C (Org.). **Manual de desenvolvimento de instrumentos psicológicos**. São Paulo: Vetor, 2017. p. 119 - 138
- COHEN, R. J.; SWERDLIK, M. E.; STURMAN, E. D. **Testagem e Avaliação Psicológica**: Introdução a testes e medidas. Porto Alegre: AMGH, 2014. 754 p.
- CONSELHO FEDERAL DE PSICOLOGIA. **Avaliação Psicológica**: Diretrizes na Regulamentação da profissão. Brasília: CFP, 2010. 196 p.
- CONSELHO FEDERAL DE PSICOLOGIA. **Cartilha Avaliação Psicológica** - 2013. Brasília: CFP, 2013. 54 p.
- CONSELHO FEDERAL DE PSICOLOGIA. Define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos e revoga a resolução CFP nº 025/2001. **Resolução CFP nº 002, de 24 de março de 2003**. Brasília: CFP, 2003
- CONSELHO FEDERAL DE PSICOLOGIA. 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. **Resolução CFP nº 9, de 25 de abril de 2018**. Brasília: CFP, 2018.
- FREIRES, L. A et al. Ensino da Avaliação Psicológica no Norte Brasileiro: analisando as ementas das disciplinas. **Avaliação Psicológica**, v. 16, n. 2, p. 2015 – 214, 2017

FREITAS, C.P. P; DAMÁSIO, B, F. Evidências de validade com base nas relações com medidas externas: conceituação e problematização. In: DAMÁSIO, B. F.; BORSA, J. C (Org.). **Manual de desenvolvimento de instrumentos psicológicos**. São Paulo: Vetor, 2017. p. 101 – 117.

GOMES, W.B. **Pesquisa e prática em Psicologia no Brasil**. Universidade Federal do Rio Grande do Sul, 2003. Disponível em: <<http://www.ufrgs.br/museupsi/ppnb.htm>>. Acesso em: 10 dez. 2017, 18:18:30

GOUVEIA, V. A Avaliação Psicológica no Brasil: Caminhos, desafios e possibilidades. **Psicologia em Foco**, Aracaju, v.2, n1, p. 110 – 119, jan./jun. 2009.

HOGAN, T.P. **Introdução á prática de testes psicológicos**. Rio de Janeiro: LTC, 2006.

MACHADO, A. P.; MORONA, V; C. **Manual de Avaliação Psicológica**. Curitiba: Unificado, 2007, 110 p.

MENDES, K.D.S; SILVEIRA, R.C.C.P; GALVÃO, C.M. **Revisão integrativa: método de pesquisa para a incorporação de evidências na saúde e na enfermagem**. Texto Contexto Enfermagem, Florianópolis, v.17, n. 4, p. 758 - 764, 2008.

MENDES – SANTOS, L.S. *et al.* Método com algoritmo específico para pontuação do teste do desenho do relógio em idosos cognitivamente normais. **Dement. Neuropsychol.**,v. 9, n.2, p. 128- 135, 2015.

MIGUEL, F. K. Padronização e interpretação de resultados. In: LINS, M.R.C.; BORSA, J. C. (Orgs.). **Avaliação Psicológica: aspectos teóricos e práticos**. Petrópolis: Vozes. 2017. p.126- 136.

NAKANO, T. C. Problemas apresentados pelos instrumentos com parecer desfavorável no SATEPSI. **Avaliação Psicológica**, Campinas, v. 12, n.2, p. 121- 130, 2013.

PACICO, J. C. Normas. In: HUTZ, C.S; BANDEIRA, D.R.; TRENTINI, C. M. **Psicometria**. Porto Alegre: Artmed, 2015, p. 45-54.

PASQUALI. Teoria da Medida. In: PASQUALI *et al.* **Instrumentação Psicológica: Fundamentos e Práticas**. Porto Alegre: Artmed, 2010, p. 55- 78.

PASQUALI, L. Validade dos testes psicológicos: Será possível reencontrar o caminho? **Psicologia: Teoria e Pesquisa**, Brasília, v. 23, n. especial, p. 099 – 107, 2007.

PRIMI, R. Avaliação Psicológica no Brasil: Fundamentos, Situação Atual e direções para o futuro. **Psicologia; Teoria e Pesquisa**. Brasília, v.26, n. especial, p.25 – 35, 2010.

PRIMI, R.C.; MUNIZ, M.; NUNES, C.H.S. Definições contemporâneas de validade de testes psicológicos. In: RUTZ, C. S. (Org.). **Avanços e polêmicas em avaliação psicológica**. São Paulo: Casa do Psicólogo, 2009. p. 243 – 265.

PRIMI, R.; NUNES, C.H.S. O SATEPSI: Desafios e propostas de aprimoramento. In: CONSELHO FEDERAL DE PSICOLOGIA (BRASIL). **Avaliação Psicológica: Diretrizes na Regulamentação da profissão**. Brasília: CFP, 2010. 196 p.

SOUZA, A. C.; ALEXANDRE, N. M. C; GUIRARDELLO, E. B. Psychometric properties in instruments evaluation of reliability and validity. **Epidemiologia e Serviços de Saúde**. Brasília, v. 26, n.3, p. 649 – 659, jul./sep. 2017.

SILVA, F. H. V. C; ALCHIERI, J. C. Avaliação Psicológica da Personalidade de Condutores: uma revisão de literatura. **Psico- USF**, São Paulo, v. 12, n.2, p. 189 – 196. 2007

URBINA, S. **Fundamentos da Testagem Psicológica**. Porto Alegre: Artmed, 2007, 320 p.

VIEIRA, R. C.; CAMPOS, R. H. Notas sobre a introdução, recepção e desenvolvimento da medida psicológica no Brasil. **Temas em Psicologia**, Belo Horizonte, v.19, n.2, p. 417 -425, 2011.

ZANON, C; FILHO, N. **Fidedignidade**. In: HUTZ, C.S; BANDEIRA, D.R.; TRENTINI, C. M. **Psicometria**. Porto Alegre: Artmed, 2015, p. 85 -95.

ZIVIANI, C.; FÉRES-CARNEIRO, T; MAGALHÃES, A. Pai e mãe na conjugalidade: Aspectos conceituais e validação de constructo. **Paideia**, Ribeirão Preto, v. 22, n. 52, p, 165 – 175, 2012.

## APÊNDICES

**APÊNDICE A – Tabela das fontes de validade referentes à palavra- chave  
Psicometria**

<b>VALIDADE</b>	<b>2017-2013</b>	<b>2012 - 2008</b>	<b>2007 – 2003</b>
Conteúdo	3	3	1
Processo de resposta	0	1	0
Estrutura interna	6	1	0
Critério	1	0	1
Preditivo	2	2	0
Concorrente	4	2	0
Convergência/ discriminante	15	3	1
Construto	3	3	1
Face	0	1	0
Aparente	0	1	0

Elaboração da autora, com base nos dados da pesquisa.



**APÊNDICE B – Tabela das fontes de validade referentes à palavra-chave validade**

<b>VALIDADE</b>	<b>2017 – 2013</b>	<b>2012 - 2008</b>	<b>2007 – 2003</b>
Conteúdo	8	4	1
Processo de resposta	0	2	0
Estrutura interna	22	5	2
Critério	8	4	1
Preditivo	4	4	1
Concorrente	13	9	8
Convergência	26	20	8
Discriminante	9	5	3
Construto	17	9	10
Nomológica	1	1	0
Aparente	0	0	1

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE C – Tabela das fontes de validade referentes à palavra-chave  
fidedignidade**

<b>VALIDADE</b>	<b>2017 – 2013</b>	<b>2012 - 2008</b>	<b>2007 – 2003</b>
Conteúdo	0	0	0
Processo de resposta	0	0	0
Estrutura interna	0	2	0
Critério	0	0	0
Preditivo	0	0	0
Concorrente	0	0	0
Convergência	0	0	0
Discriminante	0	0	0
Construto	1	1	1

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE D – Tabela das fontes de validade referentes à palavra-chave precisão**

<b>VALIDADE</b>	<b>2017 – 2013</b>	<b>2012 - 2008</b>	<b>2007 – 2003</b>
Conteúdo	0	0	0
Processo de resposta	0	0	0
Estrutura interna	0	1	0
Critério	0	1	0
Preditivo	0	0	0
Concorrente	0	0	1
Convergência	0	0	0
Discriminante	0	0	0
Construto	0	1	0
Validade de grupos (comparados)	0	0	1

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE E – Tabela das fontes de validade referentes à palavra-chave  
confiabilidade**

<b>VALIDADE</b>	<b>2017 – 2013</b>	<b>2012- 2008</b>	<b>2007 – 2003</b>
Conteúdo	0	0	0
Processo de resposta	0	0	0
Estrutura interna	0	0	0
Critério	0	0	0
Preditivo	0	0	0
Concorrente	0	0	1
Convergência	0	0	0
Discriminante	0	0	0
Construto	0	0	0

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE F - Tabela dos tipos de precisão referentes à palavra chave  
psicometria**

<b>PRECISÃO</b>	<b>2017- 2013</b>	<b>2012 – 2008</b>	<b>2007 -2003</b>
Estabilidade temporal / teste-reteste	6	0	0
Formas alternadas/ paralelas	0	0	3
Consistência interna	9	4	2
Consistência interna (alfa de cronbach)	10	7	3
Kuder – Richardson	1	0	0
Equivalência (kappa)	1	0	0
Reprodutibilidade	2	0	0

Elaboração da autora, com base nos dados da pesquisa.

**APÊNCICE G - Tabela dos tipos de precisão referentes à palavra chave validade**

<b>PRECISÃO</b>	<b>2017 – 2013</b>	<b>2012 – 2008</b>	<b>2007 – 2003</b>
Estabilidade Temporal / Teste-Reteste	3	2	3
Formas Alternadas/ Paralelas	3	0	1
Consistência Interna	27	5	8
Consistência Interna (Alfa De Cronbach)	26	24	20
Spearman Brown	1	0	0
Kuder – Richardson	0	2	0
Equivalência (Kappa)	2	1	0

Elaboração da autora, com base nos dados da pesquisa.

**APÊNCICE H - Tabela dos tipos de precisão referentes à palavra chave  
fidedignidade**

<b>PRECISÃO</b>	<b>2017 – 2013</b>	<b>2012 – 2008</b>	<b>2007 – 2003</b>
Estabilidade Temporal / Teste-Reteste	0	1	0
Formas Alternadas/ Paralelas	0	0	0
Consistência Interna	0	4	1
Consistência Interna por Alfa De Cronbach	1	2	2
Kuder – Richardson	0	0	0
Equivalência (Interavaliadores)	2	0	0

Elaboração da autora, com base nos dados da pesquisa.

**APÊNCICE I - Tabela dos tipos de precisão referentes à palavra chave precisão**

<b>PRECISÃO</b>	<b>2017 – 2013</b>	<b>2012 – 2008</b>	<b>2007 – 2003</b>
Estabilidade Temporal / Teste-Reteste	0	1	1
Formas Alternadas/ Paralelas	0	0	1
Consistência Interna	0	4	0
Consistência Interna Por Alfa De Cronbach	0	1	2
Kuder – Richardson	0	0	0
Equivalência	0	1	1

Elaboração da autora, com base nos dados da pesquisa.



**APÊNCICE J - Tabela dos tipos de precisão referentes à palavra chave  
confiabilidade**

<b>PRECISÃO</b>	<b>2017 – 2013</b>	<b>2012- 2008</b>	<b>2007 – 2003</b>
Estabilidade Temporal / Teste-Reteste	1	1	0
Formas Alternadas/ Paralelas	0	0	0
Consistência Interna	0	0	0
Consistência Interna por Alfa De Cronbach	0	1	1
Spearman Brown	0	0	0
Kuder – Richardson	0	0	0
Equivalência (Kappa)	0	0	1
Equivalência (Algoritmo de Sunderland)	1	0	0
Equivalência (CCI)	2	1	0

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE K - Relação de artigos: palavra-chave psicometria**

N	Nome	Ano	Periódico	Local
1	Escala de Identificação de Dotação e Talento: Estrutura e Consistência internas.	2017	PSICO – USF (Itatiba)	UFSCAR; UNC
2	Factorial Structure Validation of the Movement Assessment Battery for Children in School-Age Children Between 8 and 10 Years Old	2017	Paidéia (Ribeirão Preto)	UFAM, Universidade Potiguar, Brasília e SC
3	Dissatisfaction and Body Checking in Sports Scale: A New Measure for Athletes	2017	Paidéia (Ribeirão Preto)	Universidades: PE, PR, SP, MG
4	Toronto Alexithymia Scale: Adaption of Brazilian Version to Low – Educated Adults	2017	Paidéia (Ribeirão Preto)	UNIFESP
5	Revision of Criticism Avoidance dimension of Dimensional Clinical Personality Inventory	2017	Estudos de Psicologia (Campinas)	USF
6	Qualidades psicométricas do Eating Attitudes Test (EAT-26) para Adolescentes Brasileiros do Sexo Masculino	2017	Psicologia Teoria e Pesquisa (Brasília)	Universidades: PE, MG, SP
7	Validação do Function Observational Learning Questionnaire para o Português – Brasil	2017	Psicologia Teoria e Pesquisa (Brasília)	Universidade Estadual Maringá
8	Estudo Fatorial de	2017	Psico- USF	Universidades: Campinas e Paraná

	um Inventário de Práticas e Crenças Parentais		(Campinas)	
9	Children's Religious Coping Scale: Adaptation and Psychometric Properties	2017	Paideia (Ribeirão Preto)	UFRGS
10	Adaptation and Validation of Sport – Confidence Measure to Gymnasts and Football Players	2016	Psico – USF (Itatiba)	UFSC
12	Prototype Matching of Personality Disorders with the Dimensional Clinical Personality Inventory	2016	Psicologia Teoria e Pesquisa (Brasília)	USF
13	Psychometric Analysis of Disordered Eating in Sports Scale (DES)	2016	Paideia (Ribeirão Preto)	UFPE, USP, Universidade Federal de Juiz de Fora
14	Evidências de validade da Versão Brasileira da Escala Amor do Marriage and Relationship Questionnaire (MARQ)	2016	Psico – USF (Itatiba)	UFRN, PUC – RJ
15	Factor structure of Raven's Coloured Progressives Matrices	2016	Psico – USF (Itatiba)	UFSCAR, UFMG, USP – Ribeirão Preto
16	Bayley – III Scales of Infant and Toddler Development: Transcultural Adaptation and Psychometric Properties	2016	Paideia (Ribeirão Preto)	Mackenzie – SP/ FIEO – SP
17	Evaluation of the Ages and Stages Questionnaire – Brazil by Early Childhood	2016	Estudos de Psicologia (Campinas)	UFF, PUC – Rio, UERJ

	professionals			
18	Motivation Assessment Scale for Learning in Higher Education (EMAPRE – U) : Validity Evidence	2016	Psico – USF (Itatiba)	USF
19	Escala de Positividade (EP): Novas Evidências de Validade no Contexto Brasileiro	2016	Psico – USF (Itatiba)	PUC – RIO, UFRGS
20	Evidências de validade e fidedignidade da Escala Internacional de Inteligência Leiter – R para crianças dos 6 aos 8 anos	2016	Psicologia Teoria e Pesquisa (Brasília)	FIEO, USF, COIMBRA, MACKENZIE
21	Avaliação Psicométrica do Questionário de Atitudes Socioculturais em Relação à Aparência – 3 (SATAQ – 3)	2015	Psicologia Teoria e Pesquisa (Brasília)	IFMG, Universidade Cruzeiro do Sul, UFJF
22	Validity and Reability Evidence for Assessing Holland’s Career Types	2015	Paideia (Ribeirão Preto)	Universidade Federal do Recôncavo da Bahia e USF
23	A Relação entre potencial de aprendizagem e desempenho acadêmico: predição pelo teste dinâmico informatizado de raciocínio indutivo para crianças	2015	Estudos de Psicologia (Campinas)	UFSCAR, Mackenzie – SP, USF
24	Índice de comportamentos de Risco: construção e análise das propriedades psicométricas	2015	Estudos de Psicologia (Campinas)	UFRGS
25	Desenvolvimento	2015	Psicologia	Universidade Federal do Recôncavo

	perceptomotor e Escrita em crianças do Ensino Fundamental		Escolar e Educacional (Maringá)	da Bahia - Cruz das Almas – BA e USF
26	Evidências de validade de critério da BILOv3 em Crianças Gaúchas	2015	Psicologia: teoria e pesquisa (Brasília)	Universidade Federal das Ciências da Saúde de Porto Alegre Universidade Paulista UnB
27	Evidências de validade de uma Escala Destinada à Avaliação da Metacognição Infantil	2015	Psicologia: teoria e pesquisa (Brasília)	UFSCAR
28	Estudos psicométricos da Escala Eudemon de Bem –Estar Pessoal em adolescentes	2015	Psico – USF (Itatiba)	UFRGS
29	Psychometric Properties of the Brazilian 12 – item Short – from health survey version 2 (SF – 12v2)	2015	Paideia (Ribeirão Preto)	UFRJ, UFPb, UFRGS
30	Revision of dependency dimension of the dimensional clinical personality inventory	2015	Paideia (Ribeirão Preto)	USF
31	Adaptation of the sense of community index for Brazilian Children	2015	Paideia (Ribeirão Preto)	UFRGS
32	Evidências de validade da Escala de Avaliação de Relacionamento	2014	Estudos de Psicologia (Campinas)	UERJ
34	Avaliação de ansiedade e depressão em pacientes oncológicos: comparação psicométrica	2014	Psico – USF (Itatiba)	UnB
	O Cloze como instrumento de	2014	Psicologia Escolar e	Universidade Salgado de Oliveira/Universidade do Estado do

	avaliação de leitura nas séries iniciais		Educacional (Maringá)	Rio de Janeiro - RJ Universidade São Francisco – SP
35	Social skills of adolescents: convergent validity between IHSA-Del-Prette and MESSY	2014	Estudos de Psicologia (Campinas)	UFSCAR, UFMG
36	Tradução e adaptação da Escala de Potencial de Ajustamento Intercultural para a realidade brasileira	2013	Psico – USF (Itatiba)	UFU
37	Validity Evidence for the Reduced Version of the Young Parenting Inventory (YPI)	2013	Paideia (Ribeirão Preto)	Universidade Salgado de Oliveira, Niterói-RJ. UFRN e UnB
38	Exploratory Study of the Diagnostic Abilities of the Baptista Depression Scale – Adult Version (EBADEP-A)	2013	Paideia (Ribeirão Preto)	USF
39	Desenvolvimento e validação de uma escala de afetos positivos e negativos	2013	Psico – USF (Itatiba)	USF, UFRGS
40	Adaptação e validação da Escala Filadélfia de <i>Mindfulness</i> para adultos brasileiros	2012	Psico – USF (Itatiba)	UFRGS
41	Consciência metatextual: evidências de validade para instrumento de medida	2012	Psico – USF (Itatiba)	USF
42	Adaptação e validação da escala de Senso de humor situacional	2012	Psico – USF (Itatiba)	Universidade Salgado de Oliveira – Niterói
43	Pai e mãe na conjugalidade: aspectos conceituais e validação de	2012	Paideia (Ribeirão Preto)	PUC – RIO

	construto			
44	Validação brasileira do questionário de esquemas de Young: forma breve	2012	Estudos de Psicologia (Campinas)	PUC – RS
45	Evidências de validade e precisão da Escala de Coping através de Ouvir Música	2012	Psico – USF (Itatiba)	UnB, Universidade Tiradentes – Sergipe
46	Construção e estudos psicométricos de uma Escala de Avaliação da Impulsividade	2011	Psico – USF (Itatiba)	USF
47	Desenvolvimento do Instrumento de Avaliação Neuropsicológica Breve Infantil NEUPSILIN-INF	2011	Psico – USF (Itatiba)	UFRGS, PUC – RS, USP, MACKENZIE, UNIFESP
48	Evidências de validade da Escala Informatizada de Atitudes frente à Estatística – eSASPortuguês: um estudo correlacional	2011	Psico – USF (Itatiba)	USF
49	Construção e validação de instrumento para prática interventiva na adoção	2011	Paideia (Ribeirão Preto)	UFSCAR
50	Questionário de Hábitos de Estudo para estudantes universitários: validação e precisão	2011	Paideia (Ribeirão Preto)	Brasília, SP, RJ
51	Avaliação da Escala de Motivação Acadêmica em estudantes paulistas: propriedades psicométricas	2011	Psico – USF (Itatiba)	USF, Centro Universitário Adventista de SP

52	Escala de Autoestima de Rosenberg (EAR): validade fatorial e consistência interna	2010	Psico – USF (Porto Alegre)	UFRGS
53	Estudos de validade entre instrumentos que avaliam habilidades linguísticas	2010	Estudos de Psicologia (Campinas)	USF
54	Validade e confiabilidade da versão Brasileira da Center for Epidemiological Scale - Depression (CES-D) em idosos Brasileiros	2010	Psico – USF (Itatiba)	Universidades São Paulo/ Campinas/ Estados Unidos
55	Validade por processo de resposta no teste de Cloze	2009	Fractal: Revista de Psicologia (Rio de Janeiro)	USF – Itatiba
56	Inventário de Depressão Infantil (CDI): análise dos parâmetros psicométricos	2008	Fractal: Revista de Psicologia (Rio de Janeiro)	Universidades: Campinas/ Rio de Janeiro/ São Paulo
57	Construção e validação da escala de percepção de invulnerabilidade	2008	Psicologia em Estudo (Maringá)	Universidade Metodista de São Paulo
58	Propriedades psicométricas da versão brasileira da Escala de Enrincheiramento na Carreira	2008	Psico – USF (Itatiba)	Universidade Luterana do Brasil
59	Escalas de interesses vocacionais (EIV): construção, validade fatorial e consistência interna	2008	Psicologia em Estudo (Maringá)	UFRGS
60	The validity of Cloze Oriented System (COS): a correlation study with an electronic comprehension test	2007	Psicologia Escolar e Educacional (Campinas)	USF



	and a reading attitude survey			
61	Consistência interna e fatorial do Inventário Multifatorial de Coping para Adolescentes	2006	Psico – USF (Itatiba)	UFRGS Universidade do Vale do Rio dos Sinos
62	Orientação e evitação ao êxito em uma população do nordeste Brasileiro	2006	Psico – USF (Itatiba)	UNAM/ UFPb
63	Características psicométricas da Bateria Padrão do Universal Nonverbal Intelligence Test (UNIT): um estudo preliminar	2006	Psicologia Escolar e Educacional (Campinas)	UFU
64	Escala de estratégias de leitura para etapa inicial do ensino fundamental	2006	Estudos de Psicologia (Campinas)	USF
65	Estudo de validade de uma escala de desempenho em tecnologias para estudantes	2006	Psicologia Escolar e Educacional (Campinas)	USF
66	Construção e validação de uma escala de comportamentos mediacionais de educadores infantis	2006	Psicologia Escolar e Educacional (Campinas)	UFU
67	Teste de Inteligência R1- Forma B e G36: evidência de validade convergente	2005	Estudos de Psicologia (Natal)	USF
68	Evidências de validade de uma escala de desempenho docente em informática educacional	2004	Psico – USF (Itatiba)	USP e USF

69	Validação da escala de estresse no trabalho	2004	Estudos de Psicologia (Natal)	UnB
----	---	------	-------------------------------	-----

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE L - Relação de artigos: palavra-chave validade**

1	Questionário de Necessidade de Emoções (NAQ-S): Validade de Construto, Invariância e Fidedignidade	2017	Psico USF (Campinas)	UFMT, UFPb, Reino Unido
2	Posttraumatic Growth Inventory (PTGI): Adaptação e Validade Fatorial no Nordeste Brasileiro	2017	Psico USF (Campinas)	UFPI, UFPb, Universidade Federal do Vale do São Francisco
3	Evidências de Validade da Escala de Atitudes em Relação à Leitura: ERAS-Br	2017	Psico USF (Campinas)	UERJ, UFJF, Universidade Salgado de Oliveira, Rio de Janeiro
4	Validity evidence of the Zulliger-CS in older adults with Parkinson's disease	2017	Estud. psicol. (Campinas)	Universidade de Passo Fundo
5	Validity evidence of the Social and Emotional Nationwide Assessment (SENNA 1.0) Inventory	2017	Paidéia (Ribeirão Preto) vol.27	Ghent University, Ghent, Belgium e UnB
6	Alienação parental: elaboração de uma medida para mães	2017	Estud. psicol. (Campinas)	Faculdade Maurício de Nassau (PB), UFPI (Parnaíba), UFPb, Faculdades Integradas de Patos (PB),
7	Adaptation and Validation of the Brazilian DASE and TUD Scales for Cocaine/Crack Users	2017	Paidéia (Ribeirão Preto)	PUC do Rio Grande do Sul, Porto Alegre-RS, University of Maryland, USA
8	Tradução, Adaptação e Evidências Iniciais de Validade da <i>Magical Ideation Scale</i>	2017	Psic.: Teor. e Pesq. (Brasília)	USF
9	Extended Validation Study of the Thinking and Creative Style Scale: Development of a Shorter Version	2017	Psic.: Teor. e Pesq. (Brasília)	University of Ottawa PUC – Campinas
10	Escala de Autoeficácia Ocupacional em Intervenções com Populações Vulneráveis	2017	Psic.: Teor. e Pesq. (Brasília)	Universidade Salgado de Oliveira Universidade Federal do Rio

				de Janeiro Universidade Federal do Rio Grande do Sul, Optentia Research Focus Area - North- West University, Vanderbijlpark, South Africa
11	Escala de Eficácia Adaptativa: evidências de validade com base na estrutura interna e normas interpretativas	2017	Estud. psicol. (Campinas)	PUC – Campinas
12	Exploratory and confirmatory factor analysis of the <i>Roteiro de Avaliação da Consciência Fonológica</i> , a phonological awareness test	2017	Estud. psicol. (Campinas)	USF, UFB
13	Assumir o Comando no Trabalho: Evidências de Validade da Versão Brasileira da Escala	2017	Psic.: Teor. e Pesq. (Brasília)	ESSEC Business School UnB
14	Evidências de Validade de um Inventário de Competências Empreendedoras para Empresários Juniores	2017	Psico – USF (Campinas)	UFRN
15	Escala Trifatorial da Identidade Social (ETIS): Evidências de sua Adequação Psicométrica	2017	Psico – USF (Campinas)	Instituto Superior de Ciências Policiais e do Centro Universitário e UnB
16	Construção e evidência de validade do Teste de Reação à Frustração Objetivo	2017	Psic.: Teor. e Pesq. (Brasília)	UnB e USF
17	Validade do Reading the Mind in the Eyes Test em uma Amostra Brasileira	2017	Paidéia (Ribeirão Preto)	Universidade Estadual de Londrina e USF
18	Translation and Cross-Cultural Adaptation of a Brazilian Version of the Child Feeding Questionnaire	2017	Paidéia (Ribeirão Preto)	USP e Universidade de Ribeirão Preto
19	Escala de Identidade Moral: Evidências de Validade para o Contexto Brasileiro	2017	Psico – USF (Itatiba)	UnB

20	Prova de Nomeação Rápida de Figuras para Crianças: Evidências de Validade	2017	Psico – USF (Itatiba)	Universidade Estadual de Londrina e UNIFESP
21	The Experiences in Close Relationships - Relationship Structures Questionnaire (ECR-RS): validity evidence and reliability	2017	Psico – USF (Itatiba)	USP, PUC – Campinas, Universidade de Ibirapuera, Universidade Estadual do Pernambuco, Canadá
22	Escala Pessoal de Crenças no Mundo Justo: Adaptação e Evidências de Validade	2017	Psico – USF (Itatiba)	Universidade Estadual de Goiás, UnB
23	Desenvolvimento da Escala de Percepção de Futuro da Aposentadoria (EPFA) e Correlatos Psicossociais	2017	Psico – USF (Itatiba)	UFES
24	Acceptance of Dating Violence Scale: Checking its psychometric properties	2017	Psico – USF (Itatiba)	UVA, UFPb
25	A relação entre atributos, atitudes e bem-estar na mudança organizacional	2017	Psico – USF (Itatiba)	UnB
26	Evidence of factorial validity and accuracy of attitudes in mental health scale	2017	Estud. psicol. (Campinas)	UFPb, Universidade de Fortaleza
27	Development and Content validity of the CENA Program for Educational Training of the Neuropsychology of Learning, with an emphasis on executive functions and attention	2017	Dement neuropsychol. (São Paulo)	PUC – RS
28	Program of neuropsychological stimulation of cognition in students: Emphasis on executive functions - development and evidence of content validity	2017	Dement neuropsychol. (São Paulo)	PUC – RS, Mackenzie – SP, Centro Universitário Fundação Instituto de Ensino para Osasco, São Paulo
29	Social Support Scale (MOS-SSS): Analysis of the Psychometric Properties via Item Response Theory	2016	Paidéia (Ribeirão Preto)	PUC – Goiás PUC – Campinas

30	Escala de Aconselhamento Profissional e Teste de Fotos de Profissões: evidências de validade	2016	Estud. psicol. (Campinas)	USF
31	Some aspects of the validity of the Montreal Cognitive Assessment (MoCA) for evaluating cognitive impairment in Brazilian patients with Parkinson's disease	2016	Dement. neuropsychol. (SP)	USP, UNIFESP, EUA, Universidade de Ribeirão Preto
32	Análise da Estrutura Interna do Psicodiagnóstico Miocinético – PMK	2016	PSICO – USF (Itatiba)	USF E PUC – MG
33	Escala de Resistência à Mudança (RAM): Construção, Evidências Psicométricas e Versão Reduzida	2016	PSICO – USF (Itatiba)	Universidade Tecnológica Federal do Paraná Faculdade Maurício de Nassau, João Pessoa, UFPb, UFSC.
34	Team Performance: Evidence for Validity of a Measure	2016	PSICO – USF (Itatiba)	Universidade Metodista de São Paulo UnB Tribunal de Contas SP
35	Adaptação do Questionário dos Fundamentos Morais para o Português	2016	PSICO – USF (Itatiba)	UnB
36	Evidências adicionais de validade da UWES-9 em amostras brasileiras	2016	Estud. psicol. (Natal)	UFRJ, UFF, Universidade Salgado de Oliveira, UnB
37	Como Medir o Hábito? Evidências de Validade de um Índice de Autorrelato	2016	Psic.: Teor. e Pesq. (Brasília)	Centro Universitário de Brasília UnB
38	Variância Média Extraída e Confiabilidade Composta: Indicadores de Precisão	2016	Psic.: Teor. e Pesq. (Brasília)	Universidade Salgado de Oliveira UFRJ
39	Validação do Questionário de Avaliação de Habilidades Sociais, Comportamentos, Contextos para Universitários	2016	Psic.: Teor. e Pesq. (Brasília)	UESP

40	Análise psicométrica da escala de heteroavaliação de estilos de liderança	2016	Estud. psicol. (Natal)	Universidade Salgado de Oliveira UnB
41	Adaptação e evidências de validade do Recovery-Stress Questionnaire for Athletes (RESTQ-Sport) para dançarinos adolescentes (RESTQ-Dance)	2016	Estud. psicol. (Natal)	PUC – Campinas USF UFSC Universidade de Sorocaba
42	Escala de Motivação para a Leitura para Adolescentes e Jovens: Propriedades Psicométricas	2016	Psic.: Teor. e Pesq. (Brasília)	UNICAMP
43	Validity Evidence for the Turnover and Attachment Motives Survey (TAMS) in a Brazilian Sample	2016	Paidéia (Ribeirão Preto)	PUC – RS UFRS
44	Analysis of the Psychometric Properties of a Parental Alienation Scale	2016	Paidéia (Ribeirão Preto)	Universidade Tuiuti do Paraná, Curitiba
45	Zulliger (CS) in Assessing the Relational Maturity of Children	2016	Paidéia (Ribeirão Preto)	USF
46	An Ungrateful Disposition: Psychometric Properties of the Dispositional Envy Scale in Brazil	2016	Paidéia (Ribeirão Preto)	UFPI, UFMS, UFPb
47	Psychometric Analysis of Disordered Eating in Sports Scale (DES)	2016	Paidéia (Ribeirão Preto)	USP, UFJF, UFPE
48	A Teoria de Valores Refinada: associações com comportamento e evidências de validade discriminante e preditiva	2016	Psicol. USP	UnB The Hebrew University of Jerusalem Instituto Superior de Ciências Policiais
49	Escala de Expectativas de Carreira na Organização: Desenvolvimento e Evidências de Validade	2016	Psico-USF (Itatiba)	UnB
50	Behavioural Assessment of the Dysexecutive Syndrome (BADs): Adaptação e Evidências de Validade	2016	Psico-USF (Itatiba)	UFRGS

51	Evidências de Validade da Versão Brasileira da Escala Amor do <i>Marriage and Relationships Questionnaire</i> (MARQ)	2016	Psico-USF (Itatiba)	UFRN, PUC – RIO
52	Escala de Necessidade de Pertencimento: Adaptação e Evidências de Validade	2016	Psico-USF (Itatiba)	UnB
53	Validity Evidence of the Z-Test-SC for Use With Children	2016	Paidéia (Ribeirão Preto)	USF, UECE, Centro Universitário Nossa Senhora do Patrocínio, Itu-SP
54	Metamemory and aging: Psychometric properties of the Brazilian version of the Multifactorial Memory Questionnaire for elderly	2016	Dement. Neuropsychol. (SP)	USP
55	Identidade com equipes de trabalho: Teoria e medida	2016	Estud. psicol. (Natal)	Tribunal de Contas da União UnB
56	Escala de conexão com a natureza: evidências psicométricas no contexto brasileiro	2016	Estud. psicol. (Campinas)	UFPb, UFRR
57	Medida de estilos de aprendizagem para o ensino fundamental	2016	Psicol. Esc. Educ. (Maringá)	UEL, USF, Universidade Tecnológica Federal do Paraná – Apucarana
58	Self-Control, Self-Management and Entrepreneurship in Brazilian Creative Industries	2016	Paidéia (Ribeirão Preto)	UnB, UFBA, UFRN
59	Well-Being at Work Scale: Exploratory and Confirmatory Validation in the USA	2016	Paidéia (Ribeirão Preto)	UnB
60	Comparando Indicadores Psicométricos de Duas Versões Brasileiras do <i>Social Skills Rating System</i> : Uma Revisão da Literatura	2016	Psico-USF (Itatiba)	UFSCAR, Universidade Federal de São João del-Rei, UFAL
61	Mediação no Zulliger: Evidências de Validade em Amostra de Não Pacientes	2016	Psico-USF (Itatiba)	Universidade de Passo Fundo



62	Escala de Atitudes frente à Arma de Fogo (EAFAF): Evidências de Sua Adequação Psicométrica	2016	Psic.: Teor. e Pesq. (Brasília)	UFPb, Centro Universitário IESB, University of Oregon, Instituto Superior de Ciências Policiais
63	Validity and reliability of a "simplified" version of the Taylor Complex Figure Test for the assessment of older adults with low formal education	2016	Dement. neuropsychol. (SP)	UFMG
64	Escala de Autoeficácia para dirigir: construção e avaliação preliminar das propriedades psicométricas	2016	Estud. psicol. (Campinas)	USF
65	Bateria para avaliação das altas habilidades/superdotação: análise dos itens via Teoria de Resposta ao Item	2015	Estud. psicol. (Campinas)	PUC – Campinas USF
66	Estratégias de Aprendizagem no Ensino Técnico Profissional	2015	Psico-USF (Itatiba)	Universidade Tecnológica Federal do Paraná, Apucarana UEL
67	Personal and Collective Efficacy Beliefs Scales to Educators: Evidences of Validity	2015	Psico-USF (Itatiba)	UNICAMP
68	Escala de Comportamentos de <i>Bullying</i> (ECB): Elaboração e Evidências Psicométricas	2015	Psico-USF (Itatiba)	UFPI, UFPb, Universidade Federal do Vale do São Francisco, Petrolina
69	Uma medida de apego: versão brasileira da <i>Experiences in Close Relationship Scale - Reduzida</i> (ECR-R-Brasil)	2015	Psicol. USP	PUC – RIO, UFRN
70	Associação entre sinais de sofrimento psíquico até dezoito meses e rebaixamento da qualidade de vida aos seis anos de idade	2015	Psicol. USP	USP, UNIFESP
71	Evidências de Validade do	2015	Psic.: Teor. e Pesq.	UNICAMP E

	Teste Luria-Nebraska para Crianças: Relações com Escolaridade e Inteligência		(Brasília)	USP
72	Evidences of Factorial Structure and Precision of Phonemic Awareness Tasks (TCFe)	2015	Paidéia (Ribeirão Preto)	UDESC, UNIFESP
73	Evidence of Validity for the Portilho/Banas Teaching Style Questionnaire	2015	Paidéia (Ribeirão Preto)	PUC – PARANÁ UEL
74	Evidence of Validity of the Job Crafting Behaviors Scale	2015	Paidéia (Ribeirão Preto)	Universidade Salgado de Oliveira, Niterói-RJ
75	Interpersonal Interactions in the Marital Pair and Mental Health: A Comparative and Correlational Study	2015	Paidéia (Ribeirão Preto)	UFGO, PUC – GOIÁS
76	Escalas de Medida da Percepção da Qualidade do Ambiente Hospitalar - Um Estudo em Unidades de Dor	2015	Psic.: Teor. e Pesq. (Brasília)	UFPb, Instituto Universitário de Lisboa
77	Análise Fatorial Confirmatória e Normatização da <i>Hospital Anxiety and Depression Scale (HADS)</i>	2015	Psic.: Teor. e Pesq. (Brasília)	UFSE
78	Teste para Identificação de Sinais de Dislexia: processo de construção	2015	Estud. psicol. (Campinas)	PUC – CAMPINAS UNICAMP
79	Estudos psicométricos preliminares da Escala Baptista de Depressão para Adultos	2015	Estud. psicol. (Campinas)	UFF, USF, Pearson Clinical Assessment do Brasil (SP)
80	Validity Evidences for the Dimensional Clinical Personality Inventory in Outpatient Psychiatric Sample	2015	Paidéia (Ribeirão Preto)	UNIFESP, USF
81	Adaptation and Validation of the Brazilian Version of the Utrecht Work Engagement Scale	2015	Psico-USF (Itatiba)	UFRGS, Universidade de Utrecht
82	Avaliação do preconceito contra diversidade sexual e de gênero: construção de um instrumento	2015	Estud. psicol. (Campinas)	UFRGS
83	Tradução e adaptação semântica do Questionário de	2015	Estud. psicol.	PUC - RIO Centro

	Controle Atencional para o Contexto Brasileiro		(Campinas)	Universitário Augusto Motta Centro Universitário Celso Lisboa – RJ UFF, UNIFESP, PUC – RIO
84	Escala de Disposição para Perdoar: estrutura, consistência interna e invariância fatorial	2015	Estud. psicol. (Campinas)	UFPb, UFRR, Centro Universitário de João Pessoa
85	Compreensão de leitura e consciência fonológica: evidências de validade de suas medidas	2015	Estud. psicol. (Campinas)	Universidade Federal do Recôncavo da Bahia USF
86	Brazilian adaptation of the Hotel Task: a tool for the ecological assessment of executive functions	2015	Dement. neuropsychol.(SP)	PUC – RS UFPR
87	Validação da Versão Brasileira da Escala ISMI Adaptada para Dependentes de Substâncias	2015	Psic.: Teor. e Pesq. (Brasília)	UFJF, UNIFESP, University of California
88	Evidências de Validade de Construto da Escala de Componentes do Amor	2015	Psic.: Teor. e Pesq. (Brasília)	UERJ
89	Evidências de Validade de uma Escala Destinada à Avaliação da Metacognição Infantil	2015	Psic.: Teor. e Pesq. (Brasília)	UFSCAR
90	Propriedades psicométricas iniciais do Acceptance and Action Questionnaire - II - versão brasileira	2015	Psico-USF (Itatiba)	UnB
91	Teste do Desempenho Escolar: evidências de validade do subteste de escrita	2015	Psico-USF (Itatiba)	UFRGS, USF
92	Comprometimento e consentimento organizacional: um estudo da validade discriminante dos construtos	2015	Psico-USF (Itatiba)	UFBA
93	Psychometric Studies of the Learning Strategies Scale for University Students	2015	Paidéia (Ribeirão Preto)	UNICAMP E USF
94	Protocolo de Avaliação de Crianças com Autismo: Evidências de Validade de	2015	Psic.: Teor. e Pesq. (Brasília)	UFRGS

	Critério			
95	The Positivity Dimension of Well-Being: Adaptation and Psychometric Evidence of a Measure	2014	Paidéia (Ribeirão Preto)	UFPb Centro Universitário João Pessoa, João Pessoa-PB Cardiff University Cardiff, Wales
96	Relação entre medidas de Avaliação da Linguagem Escrita em estudantes do Ensino Fundamental	2014	Psico – USF (Itatiba)	Universidade Federal do Recôncavo da Bahia
97	Escala de Percepção do Suporte Social (versão adulta) - EPSUS-A: estudo das qualidades psicométricas	2014	Psico-USF (Itatiba)	USF
98	Validity evidence of the brazilian version of the five facet mindfulness questionnaire (FFMQ)	2014	Psic.: Teor. e Pesq. (Brasília)	UNIFESP, UFJF
99	Evidências de validade da Escala de Avaliação do Relacionamento	2014	Estud. psicol. (Campinas)	UERJ
100	Escala de Estratégias de Aprendizagem: evidências de validade em contexto universitário híbrido	2014	Psico-USF (Itatiba)	USP
101	Escala de Procrastinação Ativa: evidências de validade fatorial e consistência interna	2014	Psico-USF (Itatiba)	UFPb, Centro Universitário de João Pessoa, Faculdade Maurício de Nassau
102	Motivação autônoma de estudantes de física: evidências de validade de uma escala	2014	Psicol. Esc. Educ. (Maringá)	UESC, UFSC, UEL
103	An Analysis of the Factorial Structure of the Teacher Communication Behavior Questionnaire with Brazilian High School Science Students	2014	Psic.: Teor. e Pesq. (Brasília)	UFMG, Universidade de Ouro Preto, Flórida
104	Ansiedade em situações de prova: evidências de validade de duas escalas	2014	Psico-USF (Itatiba)	UnB
105	Escala de instabilidade emocional para a segurança	2014	Psico-USF (Itatiba)	Universidade Salgado de

	pública			Oliveira – RJ
106	Construção e evidências de validade de duas escalas de percepção de desenvolvimento profissional	2014	Psico-USF (Itatiba)	Universidade Salgado de Oliveira – RJ UnB
107	Evidências de validade com base na estrutura interna no Teste dos Contos de Fadas	2014	Psico-USF (Itatiba)	PUC – RS, UFSC
108	Análise Fatorial Confirmatória do Instrumento de Avaliação Neuropsicológica Breve Infantil – NEUPSILIN – Inf	2014	Psico-USF (Itatiba)	UFRGS, PUC – RS
109	Scale of Family Unpredictability During Childhood: Validity Evidence	2014	Paidéia (Ribeirão Preto)	USP, UFES
110	Escala de crenças sobre amor romântico: indicadores de validade e precisão	2014	Psic.: Teor. e Pesq. (Brasília)	UFES
111	Addenbrooke's Cognitive Examination-Revised is accurate for detecting dementia in Parkinson's disease patients with low educational level	2014	Dement. neuropsychol.(SP)	Hospital Santa Marcelina, São Paulo
112	Adaptação e evidências de validade da <i>Bateria</i> de Personalidade Prosocial no Brasi	2013	Psico-USF (Itatiba)	UnB
113	Relacionamento interpessoal, produtividade e habilidades sociais: um estudo correlacional	2013	Psico-USF (Itatiba)	Universidade de Passo Fundo
114	Evidências de validade da Escala Triangular do Amor de Sternberg – Reduzida (ETAS-R)	2013	Psico-USF (Itatiba)	UFES, Universidade Federal da Grande Dourados
115	Estudo de validação e fidedignidade de escalas de silhuetas brasileiras em adolescentes	2013	Psic.: Teor. e Pesq. (Brasília)	Universidade de Ribeirão Preto USP
116	Estudos psicométricos da escala de motivação para a aprendizagem de universitários	2013	Fractal, Rev. Psicol. (RJ)	USF E UEL
117	Evidências de validade do Experience in Close Relationships (ECR) Inventory para o Brasil	2013	Estud. psicol. (Natal)	UFRN, UFRGS

118	Intelligence assessment of deaf students with TONI 3	2013	Psico-USF (Itatiba)	Mackenzie (SP) Universidade Cruzeiro do Sul (SP)
119	Validade convergente do Inventário de Empatia (IE)	2013	Psico-USF (Itatiba)	UERJ
120	Propiedades psicométricas de anomia, alienación y desarrollo moral en estudiantes de bachillerato	2013	Psico-USF (Itatiba)	UFPb, México
121	Development of Instruments to Assess Shame and Guilt in Adolescents: Empirical Differences Between the Constructs	2013	Paidéia (Ribeirão Preto)	UFRGS
122	Estrutura fatorial do WISC-III em crianças com dificuldades de aprendizagem	2013	Psico-USF (Itatiba)	Universidade Católica de Pelotas
123	Medindo consumo de álcool: análise fatorial confirmatória do Alcohol Use Disorder Identification Test (AUDIT)	2013	Psico-USF (Itatiba)	UFC
124	Escala Diagnóstica Adaptativa Operacionalizada de Autorrelato - EDAO-AR: Evidências de Validade	2013	Paidéia (Ribeirão Preto)	PUC – Campinas
125	Validade de Critério do Inventário de Potencial para Abuso Infantil (CAP)	2013	Paidéia (Ribeirão Preto)	UFSC
126	Validity of G-36 and knowledge tests in relation to performance at a training program	2013	Estud. psicol. (Campinas)	UFMG
127	Validity and reliability of a teacher's scale developed in Brazil for assessment of hyperactive: impulsive behavior and inattention in children and adolescents	2012	Estud. psicol. (Campinas)	UNIFESP
128	Evidências de validade da Bateria Informatizada de Linguagem Oral com prova de raciocínio	2012	Estud. psicol. (Campinas)	USF
129	Evidências de Validade de Marcadores Reduzidos para a Avaliação da Personalidade no Modelo dos Cinco Grandes Fatores	2012	Psic.: Teor. e Pesq. (Brasília)	UFRGS

130	Escala Baptista de Depressão (Versão Adulto) – EBADEP-A: validade convergente e estabilidade temporal	2012	PSICO – USF (Itatiba)	Universidade Estadual Paulista, Bauru USF
131	Estudo parcial da validação do Atlas do Rorschach Sistema Compreensivo em amostra de pacientes psiquiátricos de São Paulo	2012	Psico – USF (Itatiba)	USP
132	Teste de Raciocínio Auditivo Musical (RAu): estudo inicial por meio da Teoria de Reposta ao Item	2012	Psico – USF (Itatiba)	USF, UEL
133	Montreal Battery of Evaluation of Amusia: Validity evidence and norms for adolescents in Belo Horizonte, Minas Gerais, Brazil	2012	Dement. neuropsychol. (SP)	UFMG, UFRGS
134	Maturidade perceptomotora e reconhecimento de palavras: estudo correlacional entre o Bender - Sistema de Pontuação Gradual e o Teste de Reconhecimento de Palavras	2012	Estud. psicol. (Campinas)	UFMT, USF
135	A estrutura fatorial do Teste de Criatividade Figural Infantil	2012	Psic.: Teor. e Pesq. (Brasília)	PUC – CAMPINAS USF
136	Adaptação e validação da Escala Filadélfia de <i>Mindfulness</i> para adultos brasileiros	2012	Psico-USF (Itatiba)	UFRGS
137	Consciência metatextual: evidências de validade para instrumento de medida	2012	Psico-USF (Itatiba)	USF (Itatiba)
138	Autopercepção de competências transversais de trabalho em universitários: construção de um instrumento	2012	Estud. psicol. (Natal)	Faculdade Integrada de Santa Maria UFRGS
139	Questionário de Percepção dos Pais: evidências de uma medida de estilos parentais	2012	Paidéia (Ribeirão Preto)	UnB, UFC, UFPb,
140	Adaptação e validação do Inventário de Saliência (Salience Inventory) para adultos brasileiros	2012	Paidéia (Ribeirão Preto)	UFRGS
141	A estrutura fatorial do inventário de características da personalidade	2012	Estud. psicol. (Campinas)	UFMG

142	Validade incremental do Zulliger e do Pfister no contexto da toxicomania	2012	Psico-USF	USF Université de Picardie Jules Vernes
143	Metacompreensão e inteligência: um estudo correlacional com estudantes do ensino fundamental	2012	Estud. psicol. (Natal)	USF
144	Escala de Vitalidade Subjetiva - EVS: evidências de sua adequação psicométrica	2012	Psic.: Teor. e Pesq. (Brasília)	UFPB, UFAL, Centro Universitário de João Pessoa, UFPI
145	Evidências de validade de uma versão brasileira do Inventário de Funções do Voluntariado – IFV	2011	Psico-USF (Itatiba)	UnB
146	Validade do conjunto de testes da habilidade de memória de curto-prazo (CTMC)	2011	Estud. psicol. (Natal)	UFMG
147	Validade da escala de depressão: relação com ansiedade e <i>stress</i> laboral	2011	Estud. psicol. (Campinas)	USF
148	PMK: validade preditiva do PMK em relação à presença de sintomas psicopatológicos	2011	Psico-USF (Itatiba)	UFMG
149	Propriedades psicométricas das escalas de assédio moral no trabalho - percepção e impacto	2011	Psico-USF (Itatiba)	Universidade Metodista de São Paulo UFU
150	Aceitação e rejeição entre pares e habilidades sociais em universitários	2011	Estud. psicol. (Natal)	Centro Universitário Salesiano de Americana-SP USF UESP
151	Desempenho no teste de atenção dividida como resultado da idade das pessoas	2011	Estud. psicol. (Campinas)	USF
152	Validação de conteúdo de cenas do teste de conhecimento tático no tênis	2011	Estud. psicol. (Campinas)	UFMG
153	Escala de interação com pares: construção e evidências de validade para estudantes do	2011	Psico-USF (Itatiba)	PUC – CAMPINAS Universidade



	ensino superior			Estadual de Campinas Universidade do Minho
154	Evidências de validade da versão brasileira da Escala de Autoeficácia Sexual - Função Erétil	2011	Psico-USF (Itatiba)	USF
155	Escala de exploração vocacional para estudantes de ensino médio	2011	Estud. psicol. (Campinas)	UFRGS UFSM
156	Desenvolvimento e validação preliminar de uma escala multidimensional de satisfação de vida para adolescentes	2010	Psic.: Teor. e Pesq. (Brasília)	UFRGS UFSM
157	Validade preditiva de instrumentos psicológicos usados na avaliação psicológica de condutores	2010	Psic.: Teor. e Pesq. (Brasília)	UFRN, UnB
158	Desenvolvimento de uma Lista de Verificação em Comunicação e Linguagem para os Transtornos do Espectro Autístico	2010	Psico-USF (Itatiba)	Universidade Federal do Recôncavo da Bahia UESP USP
159	Localização e qualidade formal do Rorschach-SC no Brasil: validade com não-pacientes	2010	Psico-USF (Itatiba)	USF
160	Criterion validity of a Wechsler-III scale short form in a sample of Brazilian Elderly	2010	Dement. neuropsychol. (SP)	UFJF
161	Validade e precisão do Questionário de Relacionamento Central 6.0 (CRQ 6.0) para adultos com hepatite C crônica	2010	Paidéia (Ribeirão Preto)	PUC – CAMPINAS
162	Inventário de Arnett de Busca de Sensações (AISS): testando diferentes modelos fatoriais	2010	Psico-USF (Itatiba)	UFPb, UnB, Centro Universitário de João Pessoa
163	Escala de atitudes frente à tatuagem: elaboração e evidências de validade e precisão	2010	Estud. psicol. (Campinas)	UFPb
164	Construção e validação da Escala de Crenças Parentais e Práticas de Cuidado (E-CPPC) na primeira infância	2010	Psico-USF (Itatiba)	UERJ,UFSC, USP,

165	Tradução, adaptação e exploração de propriedades psicométricas da escala de tecnoestresse (RED/TIC)	2010	Psicol. estud. (Maringá)	ULBRA
166	Brazilian caregiver version of the Apathy Scale	2009	Dement. neuropsychol. (SP)	UFMG E USP
167	Estratégias de aprendizagem e desempenho acadêmico: evidências de validade	2009	Psic.: Teor. e Pesq. (Brasília)	UEL UNICAMP USF
168	Termômetro de <i>Distress</i> : validação de um instrumento breve para avaliação diagnóstica de pacientes oncológicos	2009	Psico-USF ( Itatiba)	UnB
169	Estudo de validação e reprodutibilidade de uma escala de silhueta para adolescentes	2009	Psicol. estud. (Maringá)	USP
170	Escala de clima organizacional para organizações de saúde: desenvolvimento e estrutura fatorial	2009	Estud. psicol. (Campinas)	USF
171	Reação aos procedimentos instrucionais de um curso via <i>internet</i> : validação de uma escala	2009	Estud. psicol. (Campinas)	USP, UnB
172	Validação do Questionário de Perfis de Valores (QPV) no Brasil	2009	Psic.: Teor. e Pesq. (Brasília)	Universidade Católica de Brasília UnB
173	Validação da versão brasileira do <i>Check List</i> para Avaliação da Personalidade (PACL)	2009	Psic.: Teor. e Pesq. (Brasília)	USF
174	Evidências de validade das Matrizes Progressivas Avançadas de Raven em universitários	2009	Psico-USF (Itatiba)	Casa do Psicólogo, Itatiba, Brasil
175	The performance of the Mini-Cog in a sample of low educational level elderly	2009	Dement. neuropsychol. (SP)	UERJ
176	Versão abreviada da Escala Triangular do Amor: evidências de validade fatorial e consistência interna	2009	Estud. psicol. (Natal)	UFPb
177	Propriedades psicométricas do conjunto de testes da habilidade	2009	Psico-USF (Itatiba)	UFMG

	viso espacial			
178	Apoio social: aspectos da validade de constructo em estudantes universitários	2009	Psicol. estud. (Maringá)	UCGOIAS
179	Propriedades psicométricas do Questionário de Satisfação no Trabalho (S20/23)	2008	Psico-USF (Itatiba)	ULBRA
180	Evidências de validade entre a Escala de Depressão (EDEP), o BDI e o Inventário de Percepção de Suporte Familiar (IPSF)	2008	Psico-USF (Itatiba)	USF
181	Construção e validação da Escala de Locus de Controle Parental na Saúde	2008	Psico-USF (Itatiba)	UFMG
182	Metas de realização entre estudantes do ensino médio: evidências de validade fatorial e consistência interna de uma medida	2008	Psic.: Teor. e Pesq. (Brasília)	UFPb
183	Análise fatorial com informação completa de uma prova de compreensão em leitura em estatística	2008	Psicol. Esc. Educ. (Campinas)	USF
184	Estudo sobre as relações entre autocontrole e traços de personalidade	2008	Psicol. Esc. Educ. (Campinas)	USF
185	Estrutura fatorial da escala de atitudes frente a relacionamentos afetivos estáveis	2008	Psic.: Teor. e Pesq. (Brasília)	UFAL UFPb UnB UFPE
186	Escala de autopercepção de harter para adolescentes: um estudo de validação	2008	Psic.: Teor. e Pesq. (Brasília)	UFRGS University of Reading Universidade Federal de Ciências da Saúde de Porto Alegre
187	Propriedades psicométricas dos itens do teste WISC-III	2008	Psicol. estud. (Maringá)	Universidade Católica de Pelotas UnB
188	Versão preliminar do teste pictórico de memória: estudo de validade	2008	Estud. psicol. (Campinas)	USF
189	Construção e validação da	2008	Psicol. estud.	Universidade

	Escala de Percepção de Suporte Social		(Maringá)	Metodista de São Paulo
190	Análise da consistência interna e fatorial confirmatório do IMPRAFE-126 com praticantes de atividades físicas gaúchos	2008	Psico-USF (Itatiba)	Universidade de Sherbrooke UFRGS
191	Instrumentos de inteligência e interesses em orientação profissional	2008	Estud. psicol. (Natal)	USF
192	Escala multidimensional de satisfação de vida para crianças: estudos de construção e validação	2008	Estud. psicol. (Campinas)	UFMS, UFRGS
193	A new Brief computerized cognitive screening battery (CompCogs) for early diagnosis of Alzheimer's disease	2008	Dement. neuropsychol. (SP)	USP, UFMG, UCSP
194	Escala de auto-eficácia para atividades ocupacionais: construção e estudos exploratórios	2008	Paidéia (Ribeirão Preto)	USF
195	The systemic family assessment system: its validity with asthmatic children and their families	2007	Psicol. estud. (Maringá)	Université du Québec à Trois-Rivières UFSCAR
196	WISC-III e WAIS-III na avaliação da inteligência de cegos	2007	Psicol. estud. (Maringá)	UFMG
197	Avaliação psicológica de crianças surdas pelo Teste das Pirâmides Coloridas de Pfister	2007	Psico-USF (Itatiba)	USF
198	Estudo complementar da validade fatorial da Escala Fatorial de Satisfação em Relacionamento e predição de satisfação global com a relação	2007	Psico-USF (Itatiba)	UFSC
199	Escala de atitudes frente à escola: validade fatorial e consistência interna	2007	Psicol. Esc. Educ. (Campinas)	UFPb
200	Propriedades psicométricas do Maslach Burnout <i>Inventory</i> em uma amostra multifuncional	2007	Estud. psicol. (Campinas)	Universidade Luterana do Brasil (RS)

201	Tradução, adaptação e validação preliminar do Prospective and Retrospective Memory Questionnaire (PRMQ)	2007	Psico-USF (Itatiba)	UFRGS
202	Escala de Avaliação de Sintomas-40 (EAS-40): validade e precisão em amostra não-clínica	2007	Psicol. Esc. Educ. (Campinas)	PUC – CAMPINAS
203	Investigação das propriedades psicométricas de uma escala de metas de realização	2007	Estud. psicol. (Campinas)	Mackenzie. São Paulo USF
204	Construção e validação de uma escala de extroversão no modelo dos Cinco Grandes Fatores de Personalidade	2006	Psico-USF (Itatiba)	Faculdade Ruy Barbosa UFRGS
205	Orientação e evitação ao êxito em uma população do nordeste Brasileiro	2006	Psico-USF (Itatiba)	UNAM UFPb
206	Inteligência emocional e desempenho no trabalho: um estudo com MSCEIT, BPR-5 e 16PF	2006	Paidéia (Ribeirão Preto)	USF
207	Escala de racismo moderno: adaptação ao contexto brasileiro	2006	Psicol. estud. (Maringá)	UFPb Universidad de Almería – Espanha Universidade Tiradentes
208	Bateria de habilidades cognitivas Woodcock-Johnson III: validade de construto	2006	Psic.: Teor. e Pesq. (Brasília)	PUC – Campinas UFSCAR
209	Inteligência emocional em estudantes universitários	2006	Psic.: Teor. e Pesq. (Brasília)	USF Mackenzie
210	Escala de estratégias de leitura para etapa inicial do ensino fundamental	2006	Estud. psicol. (Campinas)	USF
211	Inteligência emocional: validade discriminante entre MSCEIT e 16 PF	2006	Paidéia (Ribeirão Preto)	USF Universidade de Alfenas
212	A construção de um instrumento de avaliação discente de um programa de pós-graduação	2005	PUC – RS	Psico-USF (Itatiba)

213	Desenvolvimento de um teste informatizado para avaliação do raciocínio, da memória e da velocidade do processamento	2005	Estud. psicol. (Campinas)	USF PUC – CAMPINAS
214	Estrutura fatorial da escala de auto-imagem: testando modelos alternativos	2005	Estud. psicol. (Campinas)	UFAL
215	Desenvolvimento do Inventário de Percepção de Suporte Familiar (IPSF): estudos psicométricos preliminares	2005	Psico-USF (Itatiba)	USF
216	Estudo de fidedignidade e validade da Escala de Avaliação de Dor Psicológica	2005	Psico-USF (Itatiba)	PUC – MG PUC – RS
217	Escala de comportamentos socialmente responsáveis do consumidor: estudo preliminar de evidência de validade	2005	Psicol. estud. (Maringá)	UNB UFPb UFRS
218	Escala sobre o clima para criatividade em sala de aula	2005	Psic.: Teor. e Pesq. (Brasília)	Universidade Católica de Brasília UNB
219	Análise fatorial do Maslach Burnout Inventory (MBI) em uma amostra de professores de instituições particulares	2004	Psicol. estud. (Maringá)	ULBRA/Canoas
220	Evidências de validade de uma escala de desempenho docente em informática educacional	2004	Psico – USF (Itatiba)	USF
221	Traços de personalidade de crianças e emoções: evidência de validade	2004	Paidéia (Ribeirão Preto)	USF
222	Análise de itens de uma prova de raciocínio estatístico	2004	Psicol. estud. (Maringá)	USF
223	Construção e validação de escala de crenças sobre o sistema treinamento	2004	Estud. psicol. (Natal)	UNB Universidade Corporativa Banco do Brasil
224	Percepção e julgamento da retaliação organizacional: construção e validação fatorial	2004	Estud. psicol. (Natal)	Universidade Católica de

	de um instrumento			Goiás UFU UNB
225	Desenvolvimento de uma escala de bem-estar subjetivo	2004	Psic.: Teor. e Pesq. (Brasília)	UNB
226	A estrutura do 16PF-5, versão espanhola: uma análise fatorial dos itens	2004	Estud. psicol. (Natal)	UFPb Universidade Complutense de Madri
227	O teste das pirâmides coloridas e o transtorno do pânico	2004	Psicol. estud. (Maringá)	USF
228	<i>Burnout e hardiness</i> : um estudo de evidência de validade	2004	Psico-USF (Itatiba)	Centro Universitário de Santo André USF
229	Estudo correlacional do inventário de busca auto dirigida (self-directed search) com o IFP	2004	Psicol. Esc. Educ. (Campinas)	USF
230	Escala de avaliação de depressão para crianças: um estudo de validação	2004	Estud. psicol. (Campinas)	PUC – CAMPINAS
231	Validação cruzada de uma escala de clima organizacional	2004	Estud. psicol. (Natal)	UNB
232	Elaboração e estudo de propriedades psicométricas do Inventário de Práticas Parentais	2003	Psico-USF (Itatiba)	Universidade do Vale do Rio dos Sinos, RS UFRGS
233	Avaliação da validade do questionário de estilo de atribuição para crianças (CASQ)	2003	Psicol. Esc. Educ. (Campinas)	UFPR
234	Adaptação da Pavlovian Temperament Survey para a realidade brasileira-versão 7 a 14 anos	2003	Psicol. Esc. Educ. (Campinas)	PUC – CAMPINAS
235	Validade da escala de estágios de mudança	2003	Estud. psicol.	PUC -

			(Campinas)	CAMPINAS USF
236	A utilização do QSG-12 na população geral: estudo de sua validade de construto	2003	Psic.: Teor. e Pesq. (Brasília)	UFPB
237	Inventário masculino dos esquemas de gênero do autoconceito (IMEGA)	2003	Psic.: Teor. e Pesq. (Brasília)	Universidade Católica de Brasília UNB
238	Competência social infantil: análise discriminante entre crianças imigrantes e não imigrantes no contexto escolar de Porto Alegre	2003	Psicol. estud. (Maringá)	PUC – RS
239	Estilos cognitivos e personalidade: um estudo exploratório de evidências de validade	2003	Psico-USF (Itatiba)	USF
240	Indicadores de alcoolismo no Teste das Pirâmides Coloridas de Max Pfister	2003	Psico-USF (Itatiba)	USF
241	Traços de personalidade na infância e distorção e integração de formas: um estudo de validade	2003	Psicol. estud. (Maringá)	USF/ MACKENZIE
242	Estilos motivacionais de professores: propriedades psicométricas de um instrumento de avaliação	2003	Psic.: Teor. e Pesq. (Brasília)	UEL Universidade Estadual de Campinas
243	Motivação do aprendiz de medicina: uso da escala de motivação acadêmica	2003	Psic.: Teor. e Pesq. (Brasília)	UnB
244	Mensuração de <i>coping</i> no ambiente ocupacional	2003	Psic.: Teor. Pesq. (Brasília)	UNB

Elaboração da autora, com base nos dados da pesquisa.



**APÊNDICE M - Relação de artigos: palavra-chave fidedignidade**

1	Psicoterapia de Crianças: Desenvolvimento da Versão em Português do <i>Child Psychotherapy Q-Set</i>	2017	Psic.: Teor. e Pesq. (Brasília)	Universidade do Vale do Rio dos Sinos San Francisco Center for Psychoanalysis
2	Methodological Criteria for Scoring Clustering and Switching in Verbal Fluency Tasks	2016	Psico – USF (Itatiba)	UFRGS
3	Propriedades psicométricas da escala de percepção de estressores ocupacionais dos professores (EPEOP)	2015	Psicol. Esc. Educ. (Maringá)	UNIFOR/ PUC – RS
4	Positive Mental Health Scale: Validation of the <i>Mental Health Continuum - Short Form</i>	2015	Psico – USF (Itatiba)	PUC – CAMPINAS UFRSG
5	Tradução, adaptação e exploração de propriedades psicométricas da Escala Interação Trabalho- Família <i>Nijmen</i> (SWING) em uma amostra de professores brasileiros	2014	Est. Psic. (Natal)	Universidade Federal de Ciências da Saúde de Porto Alegre PUC – RS
6	Motivação para mudança: análise fatorial da URICA para hábitos alimentares	2012	Psico – USF (Itatiba)	PUC – RS
7	Validação do Inventário de Concepções de Deficiência em Situações de Trabalho (ICD-ST)	2012	Psico – USF (Itatiba)	UFSJ
8	Construção de uma escala de empregabilidade: definições e variáveis psicológicas	2011	Estud. psicol. (Campinas)	USF
9	Escala de exploração vocacional para estudantes de ensino médio	2011	Estud. psicol. (Campinas)	UFRGS
10	Construção e fidedignidade teste- reteste de escalas de silhuetas brasileiras para adultos e crianças	2009	Psic.: Teor. e Pesq.(Brasília)	USP (Ribeirão Preto)
11	Reavaliando o Bem <i>Sex-Role Inventory</i>	2009	Estud. psicol. (Campinas)	Instituição Educativa São Judas Tadeu (PORTO ALEGRE)
12	Empregabilidade: Construção de uma escala	2008	Psico-USF (Itatiba)	USF

13	Avaliação estrutural da escala de ajustamento diádico	2008	Psicol. estud. (Maringá)	UFRGS
14	Mini-mac - escala de ajustamento mental para o câncer: estrutura fatorial	2008	Psicol. estud. (Maringá)	UFU
15	Inventário de Depressão de Beck - BDI: validação fatorial para mulheres com câncer	2007	Psico-USF (Itatiba)	UFU
16	Adaptação e validação de construto da Escala de Satisfação no Trabalho	2006	Psico-USF (Itatiba)	UFU
17	Estudo para a construção de uma escala de ansiedade para adolescentes	2005	Estud. psicol. (Campinas)	Universidade do Vale do Sapucaí USF
18	Fidedignidade da escala de condutas anti-sociais e delitivas ao contexto brasileiro	2003	Psicol. estud. (Maringá)	UFPb

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE N - Relação de artigos: palavra-chave precisão**

1	Escala de Aconselhamento Profissional: análise com estudantes de Ensino Médio	2012	Fractal (RJ)	USF UFRGS
2	Estudo nacional dos atributos psicométricos da Escala Triangular do Amor de Sternberg	2012	Paideia (Ribeirão Preto)	Universidade Federal da Grande Dourados UnB
3	Evidências desfavoráveis para avaliação da personalidade com um instrumento de 10 itens	2012	Paideia (Ribeirão Preto)	Mackenzie – SP UFRGS UFSC USF
4	Características psicométricas da <i>Relationship Assessment Scale</i>	2011	Psico – USF (Itatiba)	Universidade Federal da Grande Dourados UNB
5	Escala Baptista de Depressão (Versão Adulto) - EBADEP-A: evidências de validade de construto e de critério	2011	Psico – USF (Itatiba)	USF
6	Adequação ao Real de adolescentes: possibilidades informativas do Questionário Desiderativo	2009	Psic.: Teor. e Pesq. (Brasília)	Universidade de São Paulo (São José do Rio Preto)
7	Comparação entre dois sistemas de pontuação para o teste informatizado de percepção de emoções em fotos	2009	Estud. psicol. (Campinas)	USF
8	Escala de Preferência Musical: construção e comprovação da sua estrutura fatorial	2007	PSICO –USF (Itatiba)	UFPb
9	Precisão de avaliadores na avaliação da criatividade por meio da produção de metáforas	2007	PSICO –USF (Itatiba)	USF
10	Escala de Avaliação de Sintomas-40 (EAS-40): validade e precisão em amostra não-clínica	2007	Psicol. Esc. Educ. (Campinas)	PUC – CAMPINAS
11	Bateria Multidimensional de Inteligência Infantil: desenvolvimento de instrumento	2005	PSICO –USF (Itatiba)	UFSCAR, PUC – CAMPINAS
12	Avaliação preliminar da escala de desempenho em informática educacional com professores	2004	Estud. psicol. (Campinas)	USF
13	Atitudes de estudantes	2004	Estud. psicol.	UFPb

	universitários frente ao consumo de materiais pornográficos		(Natal)	
14	Desenvolvimento e validação da escala de civismo nas organizações	2003	Estud. psicol. (Natal)	UNB

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE O - Relação de artigos: palavra-chave normas**

1	Normative data for the Brief Cognitive Screening Battery stratified by age and education	2017	Dement. neuropsychol. (SP)	Universidade Federal do ABC; USP; UNICAMP
2	Influence of age and education on the Rivermead Behavioral Memory Test (RBMT) among healthy elderly	2016	Dement. neuropsychol. (SP)	Universidade Católica do Rio Grande do Sul UFRGS USP PUC – RS
3	Rorschach: Normas para Adolescentes em Diferentes Etapas da Vida	2015	Psic.: Teor. e Pesq. (Brasília)	Puc –SP
4	Normative Study of Rorschach (Parisian School) for Brazilian Adolescents	2015	Paidéia (Ribeirão Preto)	USP – Ribeirão Preto
5	Brazilian preliminary norms and investigation of age and education effects on the Modified Wisconsin Card Sorting Test, Stroop Color and Word test and Digit Span test in adults	2015	Dement. neuropsychol. (SP)	UNIFESP PUC – RS UFRGS PUC – RS
6	Mattis Dementia Rating Scale (DRS): Normative data for the Brazilian middle-age and elderly populations	2013	Dement. neuropsychol. (SP)	USP – RP UFMG FMRP-USP USP
7	The influence of schooling on performance in the Mattis Dementia Rating Scale (DRS)	2010	Dement. neuropsychol. (SP)	USP UFMG
8	Normas de emocionalidade para a versão brasileira do paradigma Deese-Roediger-McDermott (DRM)	2009	Psic.: Teor. e Pesq. (Brasília)	UFRR UFRB PUC – RS
9	Typical performance of elderly patients with Alzheimer disease on the Wisconsin Card Sorting Test (WCST)	2007	Dement. neuropsychol. (SP)	USP UFAM FMRP-USP
10	Matrizes progressivas coloridas de Raven — escala especial: normas para Porto Alegre, RS	2004	Psicol. estud. (Maringá)	USP UFRGS

11	PTS - Pavlovian Temperament Survey, versão adolescente/adulto: consistência interna e normatização para a realidade brasileira	2003	Estud. psicol. (Natal)	PUC – CAMPINAS USF
----	--	------	------------------------	-----------------------

Elaboração da autora, com base nos dados da pesquisa.

**APÊNDICE P - Relação de artigos: palavra-chave confiabilidade**

1	Specific algorithm method of scoring the Clock Drawing Test applied in cognitively normal elderly	2015	Dement. neuropsychol. (SP)	PUC – RIO King's College London, UK
2	Accuracy and reliability of the Pfeffer Questionnaire for the Brazilian elderly population	2015	Dement. neuropsychol. (SP)	UCB
3	Reliability of Cognitive Tests of ELSA-Brasil, the Brazilian Longitudinal Study of Adult Health	2013	Dement. neuropsychol. (SP)	UFMG, UFOP
4	Cross-cultural Adaptation of the "Functional Activities Questionnaire - FAQ" for use in Brazil	2011	Dement. neuropsychol. (SP)	UERJ
5	Brief cognitive battery in the diagnosis of mild Alzheimer's disease in subjects with medium and high levels of education	2007	Dement. neuropsychol. (SP)	USP UFMG
6	Escala de afeto positivo e negativo para crianças: estudos de construção e validação	2006	Psicol. Esc. Educ. (Campinas)	UFSM UFRGS

Elaboração da autora, com base nos dados da pesquisa.