

Diretrizes para a Construção de Testes Psicológicos: a Resolução CFP nº 009/2018 em Destaque

Josemberg Moura de Andrade¹
¹Universidade de Brasília, DF, Brasil.

Felipe Valentini²
²Universidade São Francisco, SP, Brasil.

Resumo: O presente artigo objetiva apresentar diretrizes para a construção de testes psicológicos considerando as recomendações do Sistema de Avaliação de Testes Psicológicos – Satepsi, a partir da Resolução do CFP nº 009, de 25 de abril de 2018. São discutidos os principais aspectos psicométricos considerados indispensáveis para a construção de testes psicológicos, a saber: evidências de precisão/fidedignidade, evidências de validade, e, por fim, sistema de correção e interpretação dos escores. Em cada uma das subseções são apresentadas brevemente as definições e formas de obtenção dos respectivos parâmetros. Além disso, destaque foi dado às modificações apresentadas pela Resolução nº 009/2018. Para além dos conceitos psicométricos, a referida Resolução também destaca a justiça e proteção dos direitos humanos no processo de avaliação psicológica. Argumenta-se que a justiça é uma questão fundamental de validade da interpretação dos escores individuais para os usos pretendidos e requer atenção em todos os estágios de desenvolvimento e uso de testes. Conclui-se que a nova Resolução apresenta avanços em relação as Resoluções e notas técnicas anteriores. A partir de um contínuo aprimoramento, o Satepsi buscou critérios psicométricos alinhados aos avanços técnicos e científicos da área. Com isso, esperamos que a Resolução impulse os desenvolvedores de testes a buscarem o aprimoramento dos testes e, conseqüentemente, do processo de avaliação psicológica.

Palavras-chave: Testes Psicológicos, Psicometria, Validade, Fidedignidade, Normas.

Guidelines for the Construction of Psychological Tests: Regulation CFP No: 009/2018 in Highlight

Abstract: This paper is dedicated to present the guidelines on constructing psychological tests based on the Satepsi (a Brazilian evaluation system for psychological tests), from CFP's (Brazilian National Council for Psychologists) Resolution No: 009/2018. We present the main psychometric features as references for constructing tests, as follow: evidence of reliability, and validity, as well as score standardization. In each section, we briefly describe the definition and the ways to assess these features. Furthermore, we highlighted the modifications yielded by the new legislation No: 009/2018. Beyond the psychometrics, the CFP's legislation also emphasizes the fairness and the human rights on assessment. We consider the fairness as one of the main aspects for scores interpretation, and it requires attention on all steps in using a test. We conclude the new CFP's legislation is more updated than the previous one. Based on a permanent attempt to enrich the standards for testing, the Satepsi seeks for psychometric criteria aligned with the technical and scientific improvements. We expect the CFP's legislation propels the test developers to constantly pursue improvements on the test and psychological assessment.

Keywords: Psychological Tests, Psychometric, Validity, Reliability, Standards.

Directrices para la Construcción de Pruebas Psicológicas: la Resolución CFP nº 009/2018 en Destaque

Resumen: El presente artículo tiene como objetivo presentar directrices para la construcción de pruebas psicológicas considerando las recomendaciones del Sistema de Evaluación de Pruebas Psicológicas - Satepsi, a partir de la Resolución del CFP nº 009, del 25 de abril de 2018. Se discuten los principales aspectos psicométricos considerados indispensables para la construcción de pruebas psicológicas, a saber: evidencias de precisión/fiabilidad, evidencias de validez, y, por último, sistema de corrección e interpretación de los escores. En cada una de las subsecciones se presentan brevemente las definiciones y formas de obtención de sus parámetros. Además, se destacó la modificación presentada por la Resolución 009/2018. Además de los conceptos psicométricos, la referida Resolución también destaca la justicia y protección de los derechos humanos en el proceso de evaluación psicológica. Se argumenta que la justicia es una cuestión fundamental de validez de la interpretación de los escores individuales para los usos pretendidos y requiere atención en todas las etapas de desarrollo y uso de pruebas. Se concluye que la nueva Resolución presenta avances en relación a las Resoluciones y notas técnicas anteriores. A partir de un continuo perfeccionamiento, el Satepsi buscó criterios psicométricos alineados a los avances técnicos y científicos del área. Con eso, esperamos que la Resolución impulse a los desarrolladores de pruebas a buscar el perfeccionamiento de las pruebas y, consecuentemente, del proceso de evaluación psicológica.

Palabras clave: Pruebas Psicológicas, Psicometría, Validez, Fiabilidad, Normas.

A avaliação psicológica e a avaliação educacional estão entre as contribuições mais importantes das ciências cognitivas para nossa sociedade. Tais avaliações fornecem informações fundamentais para a tomada de decisão e impactam na vida dos indivíduos, grupos e na sociedade como um todo. Entre os atores impactados pelas avaliações podemos citar, os estudantes, pais, familiares, professores, administradores educacionais, candidatos a vagas de emprego, entre outros. As instituições impactadas pelas avaliações, por sua vez, incluem escolas, universidades, empresas, agências governamentais etc. As instituições e os indivíduos são beneficiados quando a avaliação os orientam ao alcance dos seus objetivos. A sociedade, por sua vez, beneficia-se quando a avaliação contribui para a realização individual e para os objetivos das instituições (AERA, APA, & NCME, 2014). A partir das informações provenientes de tais avaliações, planos de intervenção individuais podem ser operacionalizados e políticas públicas podem ser delineadas no contexto macrossocial (Bauer, Alavarse, & Oliveira, 2015). Especificamente, a avaliação psicológica traz inúmeros benefícios para a sociedade. É reconhecido o potencial da área para construir

conhecimentos a respeito de construtos psicológicos, bem como para produzir, orientar, monitorar e encaminhar ações e intervenções sobre os indivíduos avaliados (Reppold, 2011).

Considerando que a testagem é um dos recursos utilizados nas avaliações, ressalta-se, no entanto, que nem todos os testes são bem construídos, nem todas as práticas de testagem são adequadas ou benéficas. Há evidências que documentam a utilidade de testes quando bem construídos e bem interpretados. Testes bem desenvolvidos e que são válidos para seus propósitos pretendidos têm potencial de oferecer benefícios efetivos para avaliadores e avaliados. O uso adequado de tais testes pode resultar em melhores decisões para os indivíduos do que resultaria sem seu uso. Por outro lado, o uso inapropriado pode causar danos irreparáveis para os avaliados e outros atores afetados por decisões baseadas em testagens (AERA et al., 2014).

Considerando a importância da testagem enquanto etapa do processo de avaliação psicológica, o presente artigo objetiva apresentar diretrizes para a construção de testes psicológicos. Para isso, consideramos as recomendações do Sistema de Avaliação de Testes Psicológicos – Satepsi, a partir da Resolução do

Conselho Federal de Psicologia (CFP) nº 009, de 25 de abril de 2018. São discutidos os principais parâmetros psicométricos considerados indispensáveis para a construção de testes psicológicos, a saber: evidências de precisão/fidedignidade, evidências de validade, e, por fim, sistema de correção e interpretação dos escores. O artigo tem a intenção de apresentar os conceitos psicométricos de forma clara e simples, a fim de favorecer a compreensão dos leitores menos experientes na área, no entanto, sem perder a tecnicidade.

Entendemos que o conhecimento em psicometria é de suma importância não apenas para o(a) profissional que elabora os testes psicológicos, mas, sobretudo, para o(a) psicólogo(a) que atua na prática da área da avaliação psicológica. A psicometria trata basicamente da teoria e técnica de medida dos processos mentais, especialmente aplicada nas áreas da Psicologia e Educação (Pasquali, 2013). A partir do conhecimento em psicometria, o psicólogo avaliador poderá escolher o(s) teste(s) psicológico(s) que fará(ão) parte do processo de avaliação psicológica de forma mais consciente e integrada com as demais técnicas utilizadas. Tal afirmativa encontra sustentação no Art. 2 da Resolução nº 009/2018 no qual destaca que na realização da avaliação psicológica, o psicólogo deve basear sua decisão, obrigatoriamente, em métodos e/ou técnicas e/ou instrumentos psicológicos reconhecidos cientificamente para uso na prática profissional, ou seja, as fontes fundamentais de informação. A depender do contexto, o psicólogo também poderá recorrer a procedimentos e recursos auxiliares, identificados como fontes complementares de informação. A propósito, a diferenciação entre técnicas fundamentais e técnicas complementares de informação é uma novidade na referida Resolução (009/2018).

A Resolução CFP nº 009/2018 estabelece (I) diretrizes para a realização de Avaliação Psicológica no exercício profissional dos(as) psicólogos(as), (II) regulamenta o Satepsi e (III) revoga as Resoluções nº 002/2003, nº 006/2004 e nº 005/2012, bem como as Notas Técnicas nº 01/2017 e 02/2017. O Satepsi, por sua vez, é um sistema informatizado que tem por objetivo avaliar a qualidade técnico-científica de instrumentos submetidos à apreciação da Comissão Consultiva em Avaliação Psicológica do CFP. É importante assinalar que a Resolução CFP nº 009/2018 apresenta avanços em relação às respectivas Resoluções e Notas Técnicas anteriores na medida em que

especifica mais detalhadamente os critérios mínimos considerados adequados na elaboração e obtenção de evidências de validade e fidedignidade de um teste psicológico. A partir de uma constante consulta às diretrizes (*Guidelines*) do *International Test Commission*, bem como aos *Standards for Educational and Psychological Testing* (AERA et al., 2014), buscou-se critérios psicométricos alinhados aos avanços técnicos científicos da área, considerando, no entanto, que a área possui muitas especificidades. Algumas áreas de aplicação da avaliação psicológica estão em amplo desenvolvimento, enquanto outras áreas como, por exemplo, avaliação psicológica de pessoas com necessidades especiais e pessoas com baixo nível educacional, ainda são incipientes. Destaca-se também que a Resolução CFP nº 009/2018 é fruto de uma ampla discussão com as entidades brasileiras da área da avaliação psicológica, bem como resultante do constante trabalho do CFP de aprimoramento e incorporação de melhorias do Satepsi sugeridas e debatidas em diferentes fóruns científicos.

A Resolução nº CFP 009/2018 também destaca a submissão ao Satepsi de versões equivalentes de testes psicológicos aprovados (informatizados e não informatizados), bem como a atualização das normas e dos estudos de evidências de validade dos testes psicológicos. Em relação ao primeiro aspecto, de acordo com o Art. 18, é considerada versão equivalente de um teste psicológico aquela com formato diferente de aplicação descrita na versão inicial do teste aprovado pelo Satepsi. Em relação ao segundo aspecto – atualização de normas e estudos de validade –, no Art. 14 é assinalado que os estudos de evidências de validade, precisão e normas dos testes psicológicos terão prazo máximo de 15 anos, a contar da data da aprovação do teste psicológico pela Plenária do CFP. Anteriormente, na Resolução nº 002/2003, alterada pelas Resoluções nº 006/2004 e Resolução CFP nº 005/2012, os dados empíricos das propriedades dos testes deveriam ser revisados periodicamente, não podendo o intervalo entre um estudo e outro ultrapassar 15 anos para os dados referentes à padronização, e 20 anos para os dados referentes à validade e precisão.

Para além dos conceitos psicométricos, a referida Resolução também destaca a justiça e proteção dos direitos humanos no processo de avaliação psicológica. Por exemplo, no Art. nº 31 é assinalado que é vedado a(o) psicólogo(a): a) realizar atividades que caracterizem negligência, preconceito, exploração,

violência, crueldade ou opressão; b) induzir a convicções políticas, filosóficas, morais, ideológicas, religiosas, raciais, de orientação sexual e identidade de gênero; e c) favorecer o uso de conhecimento da ciência psicológica e normatizar a utilização de práticas psicológicas como instrumentos de castigo, tortura ou qualquer forma de violência.

O Art. 31 da Resolução nº 009/2018 apresenta-se alinhado ao capítulo intitulado de “*Fairness in Testing*” (tradução livre de “Justiça na Testagem”) do *Standards for Educational and Psychological Testing* (AERA et al., 2014). Tal capítulo destaca a importância da justiça como uma questão fundamental na proteção dos direitos dos testandos em todos os aspectos e momentos da testagem. Ressalta-se que o termo justiça não tem um significado técnico único e pode ser utilizado de diferentes maneiras. A justiça é uma questão fundamental de validade da interpretação dos escores individuais para os usos pretendidos e requer atenção em todos os estágios de desenvolvimento e uso de testes.

Diante do exposto e considerando o objetivo de apresentar diretrizes para a construção de testes psicológicos a partir da Resolução nº 009/2018, os seguintes parâmetros psicométricos são discutidos na próxima seção: (I) Evidências de precisão/fidedignidade, (II) Evidências de validade; e, por fim, (III) Sistema de correção e interpretação dos escores. Em cada uma das subseções são apresentadas brevemente as definições e formas de obtenção dos respectivos parâmetros. Além disso, amplo destaque foi dado as modificações que a Resolução nº 009/2018 apresentou em relação às Resoluções anteriores nº 002/2003, nº 006/2004 e nº 005/2012.

Parâmetros psicométricos

De acordo com o Art. 1 da Resolução nº 009/2018, a avaliação psicológica pode ser definida como um processo estruturado de investigação de fenômenos psicológicos, composto de métodos, técnicas e instrumentos, com o objetivo de fornecer informações para a tomada de decisão, no âmbito individual, grupal ou institucional. Assim, o teste psicológico é um dos recursos que pode ser utilizado no processo mais amplo de avaliação psicológica (Andrade, & Sales, 2017). No Art. 4 da Resolução (009/2018), encontramos que um teste psicológico, por sua vez, tem por objetivo identificar, descrever, qualificar e mensurar características psicológicas, por meio de procedimen-

tos sistemáticos de observação e descrição do comportamento humano, nas suas diversas formas de expressão. Destacamos que na Resolução foi adotada a nomenclatura de teste para abarcar também escalas, inventários, questionários e métodos projetivos/expressivos. Importante destacar que existem diversas classificações dos testes psicológicos. Por exemplo, de acordo com os *Standards for Educational and Psychological Testing* (AERA et al., 2014), existem seis tipos de categorias de testes, a saber: (I) cognitivos e neuropsicológicos; (II) Família e casal; (III) problemas comportamentais; (IV) comportamento social e adaptativo; (V) personalidade; e (VI) vocacional.

Considerando todas as possibilidades de construtos avaliados, no Art. 6 da Resolução nº 009/2018 é assinalado que os testes psicológicos, para serem reconhecidos como técnica fundamental de informação para o uso profissional, devem possuir consistência técnico-científica. Os requisitos mínimos obrigatórios que os construtores de testes psicológicos devem atender são: I - apresentação de fundamentação teórica, com especial ênfase na definição do(s) construto(s), descrevendo seus aspectos constitutivo e operacional; II - definição dos objetivos do teste e contexto de aplicação, detalhando a população-alvo; III - pertinência teórica e qualidade técnica dos estímulos utilizados nos testes; IV - apresentação de evidências empíricas sobre as características técnicas dos itens do teste, exceto para os métodos projetivos/expressivos; V - apresentação de evidências empíricas de validade e estimativas de precisão das interpretações para os resultados do teste, caracterizando os procedimentos e os critérios adotados na investigação; VI - apresentação do sistema de correção e interpretação dos escores, explicitando a lógica que fundamenta o procedimento, em função do sistema de interpretação adotado; VII - apresentação explícita da aplicação e correção para que haja a garantia da uniformidade dos procedimentos.

Ressaltamos que todos os requisitos mínimos citados no Art. 6 são de extrema importância para a elaboração e comprovação da adequação de um teste. A propósito consideramos que essas etapas são inter-relacionadas. Por exemplo, Pasquali (2013) assinala que a delimitação teórica dos construtos continua sendo um ponto muito deficiente na Psicologia. A elaboração das definições constitutivas e operacionais, a partir de um marco teórico específico, terá um impacto direto na qualidade do teste que se quer construir.

Sem minimizar a importância de todos os requisitos, o presente artigo focalizou no polo analítico da construção e obtenção de evidências de validade, precisão e normas dos testes. Esses requisitos são tratados nas subseções a seguir. Optamos por não discutir em uma subseção específica o tópico B.5 do Anexo da Resolução nº 009/2018, referente à Análise de itens, simplesmente devido as limitações de espaço. Ressaltamos que a análise dos itens é igualmente importante e, inclusive, a nova Resolução (009/2018) traz mais claramente quais são os critérios esperados para essa análise.

Na análise dos itens, os parâmetros psicométricos indicam propriedades (características) estatísticas desejáveis dos itens (incluindo: dificuldade, discriminação e correlações inter-item), bem como as propriedades estatísticas desejadas do teste como um todo. Neste último citamos os índices de dificuldade e precisão do teste como um todo, além da distribuição do quantitativo de itens por dimensão/fator avaliado. Quando os índices dos itens são estimados a partir da Teoria de Resposta ao Item (TRI), o ajuste do modelo aos dados precisa ser avaliado (AERA et al., 2014). Na Resolução nº 009/2018 para que o teste seja avaliado como excelente (Nível A+) é necessário conter no manual pelo mínimo três estudos distintos satisfatórios. Para o teste ser considerado suficiente (Nível B) deve conter pelo menos um estudo satisfatório. A seguir são apresentadas as subseções que tratam das evidências de validade, precisão e normas dos testes.

I – Evidências de precisão/fidedignidade

Um teste psicológico é sempre uma “ponte” entre um conjunto de comportamentos observados e o construto para o qual o instrumento se destina. No entanto, essa “ponte” entre o construto psicológico e o comportamento observado sofre influência de erros, que são variações aleatórias em relação à medida. Quanto maior for esse erro (de medida), menor é a precisão dos escores. Nesse sentido, parece bastante relevante considerar o erro de medida na construção e na utilização de um teste (AERA et al., 2014). De maneira que um diagnóstico realizado com base em instrumentos cujos escores são pouco precisos terá maiores chances de erro.

Ressalta-se que diferentes fontes de erro podem afetar os escores dos indivíduos (Nunnally, 1994; Thompson, 2002). Entre as mais diversas possibilidades de interferências aleatórias nos escores, a Resolução nº 009/2018 do CFP destaca, em seu anexo, as

condições de testagem do examinando, o formato do instrumento e as diferenças entre os avaliadores. Tais fontes, segundo os *Standards* (AERA et al., 2014) podem ser reunidas em dois grandes grupos: próprio sujeito e externas a ele.

No que se refere ao examinando, flutuações gerais nas condições físicas e psicológicas, não relacionadas ao construto psicológico mensurado podem diminuir a precisão dos escores. Por exemplo, um sujeito que teve as suas habilidades cognitivas avaliadas logo após receber uma notícia ruim pode ter o seu escore subestimado. Nesse sentido, diferentes níveis pessoais de motivação, atenção, engajamento na tarefa, fadiga, entre outros, podem inserir flutuações nos escores dos testes, o que aumenta o erro de medida devido às condições do examinando.

No que se refere às fontes externas ao avaliando, as flutuações podem ser decorrentes das condições de testagem (tais como diferenças de luminosidade, conforto do mobiliário e temperatura) e das diferenças entre avaliadores. Neste último, os aspectos subjetivos dos avaliadores (tais como atenção às respostas do examinando e rigidez ou flexibilidade no uso dos critérios de pontuação), se variados aleatoriamente, podem inserir imprecisão na estimação dos escores.

Além das fontes externas e internas, o próprio formato e conteúdo dos itens pode ser fonte de erro de medida. Nesse contexto, espera-se que o conjunto de itens seja o mais consistente entre si, refletindo os aspectos psicológicos do indivíduo e não variações aleatórias a situações específicas. Por exemplo, no intuito de avaliar a Extroversão, o psicólogo resolve construir um instrumento de autorrelato com os seguintes itens: “gosto de ir a bares com os amigos”, “gosto de ir para a balada” e “me sinto bem quando estou com amigos”. Nesse exemplo, é possível que os dois primeiros itens captem variações específicas de gostos por festas e bares, além da “energia para relações sociais”, característica de pessoas extrovertidas. No entanto, é possível que uma pessoa goste de ter muitos amigos, tenha bastante motivação e habilidades nas relações sociais, enfim, seja extrovertido, mas não goste especificamente de bares e baladas.

Ainda no que se refere ao conteúdo dos testes, diferenças muito grandes entre as características do examinando (quanto ao traço psicológico) e o nível de exigência do item também podem gerar imprecisão nos escores. Questões de raciocínio abstrato muito fáceis podem não gerar muita informação precisa

sobre examinandos com altas habilidades cognitivas, por exemplo. O contrário também é verdadeiro. Assim, a inespecificidade dos itens para o nível do traço psicológico do sujeito também pode ser fonte de erro de medida.

Obviamente, não é possível controlar todas as fontes de erro de medida, mas é possível estimar, em certo grau, a sua influência nos escores. Tais estimativas são obtidas por meio dos coeficientes de precisão (ou de informação na TRI). Quanto maior for o valor desse coeficiente, menor é o erro de medida (ou seja, as flutuações aleatórias nos escores são menores, e a precisão é maior). Cada coeficiente é obtido por um método específico e controla um tipo de erro (AERA et al., 2014).

As flutuações relacionadas as condições internas do examinando podem ser estimadas por teste-reteste. Seleciona-se uma amostra e se aplica o mesmo instrumento (ou versões equivalentes) em dois ou mais momentos distintos, com um intervalo de tempo suficiente para minimizar efeitos de memória (Nunnally, 1994). Neste caso, assume-se que a primeira aplicação não tem influência na segunda. Esse método é mais apropriado para construtos que sabidamente não sofrem alterações substanciais em curso espaço de tempo (uma ou duas semanas), como por exemplo, personalidade e inteligência. Por outro lado, se houver variações sistemáticas do construto psicológico em curso espaço de tempo, como ansiedade de estado, por exemplo, não será possível separar as flutuações genéricas aleatórias (que são erros de medida) das flutuações reais do construto, e o método de teste-reteste se torna ineficiente.

As flutuações relacionadas às condições externas ao examinando podem ser avaliadas por diferentes métodos. Caso a fonte de erro seja ambiental (por exemplo, condições de luminosidade ou temperatura), as flutuações podem ser estimadas por meio de teste-reteste variando as condições do ambiente. Se a pontuação das respostas dos indivíduos depender do juízo do profissional de Psicologia, é possível estimar a consistência (ou concordância) entre avaliadores. Neste caso, são comparadas as avaliações do mesmo sujeito realizadas por diferentes profissionais. Quanto maior for a concordância entre os avaliadores, menor é a influência de variáveis subjetivas do avaliador.

Finalmente, flutuações devido ao conteúdo específico dos itens podem ser estimadas por meio de consistência interna. Para tanto o método mais utilizado é calcular o coeficiente Alfa dos itens, após a aplica-

ção do instrumento em uma amostra de examinandos. Contudo, existem outras possibilidades como o Kuder-Richardson, mais apropriado a itens dicotômicos. Ressalta-se que esses métodos mais tradicionais assumem que os itens são igualmente difíceis e com mesmo nível de discriminação (Fornell, & Larker, 1981; Valentini, & Damásio, 2016). Esses pressupostos (muito restritivos) não são assumidos por modelagens latentes, nas quais os parâmetros de dificuldade e discriminação podem ser estimados simultaneamente ao erro de medida. No contexto da modelagem por equações estruturais, é possível estimar a confiabilidade composta, a variância média extraída (Fornell, & Larker, 1981; Valentini, & Damásio, 2016) e ômega hierárquico (Rios, & Wells, 2014).

Contudo, mesmo os métodos de estimação da consistência interna em equações estruturais estimam apenas um coeficiente, de maneira independente ao nível do construto psicológico do examinando (ressalta-se que isso não se aplica a todos os métodos de estimação dos parâmetros, pois algumas configurações em equações estruturais são equivalentes à TRI, veja, por exemplo, Kamata, & Bauer, 2008). Assim, uma das vantagens principais da TRI é oferecer uma estimativa de erro do escore de cada indivíduo, que pondera a distância entre o nível do construto psicológico e a dificuldade do item. Neste caso, a precisão será maior para os escores dos examinandos próximos do nível de dificuldade dos itens; e, por outro lado, os escores serão muito imprecisos caso o nível de exigência dos itens seja muito superior ou muito inferior ao nível de construto psicológico do examinando (Andrade, Laros, & Gouveia, 2010; Hambleton, & Swaminathan, 1991; Valentini, & Laros, 2011).

Os aspectos de precisão abarcados na Resolução nº 009/2018 do CFP diferem da nº 002/2003 principalmente sobre o espaço para esses indicadores de precisão no contexto de modelagens latentes (equações estruturais e TRI). O uso desse tipo de modelagem se popularizou consideravelmente nos últimos 15 anos, e diversos instrumentos novos já apresentam indicadores de precisão com base em TRI. Nesse sentido, é justificável e importante o espaço adicional da nova resolução dado à avaliação da precisão dos escores latentes, seja por TRI ou qualquer outra modelagem. No entanto, o conteúdo técnico desse tipo de indicador ainda não é dominado por todos os usuários profissionais de testes psicológicos. Para minimizar esse problema, sugerimos que os autores dediquem espaço

no manual dos testes para explicar de maneira simples o funcionamento desses indicadores e como utilizá-los na prática profissional de avaliação psicológica.

Destaca-se que a nova Resolução do CFP (009/2018) mantém a avaliação dos indicadores de precisão descritos no manual entre os níveis C (insuficiente) e A+ (Excelente). Para alcançar o nível de excelência exige-se dois ou mais estudos satisfatórios, com indicadores superiores a 0,80. No entanto, para ser aprovado, os escores devem apresentar indicadores de precisão acima de 0,60. Justifica-se essa diferença entre o mínimo exigido e o de excelência em função da variabilidade da importância da precisão. Quanto mais importante for a avaliação na vida do examinando e mais central for o construto investigado pelo teste, maior será a necessidade de precisão (AERA et al., 2014). Por exemplo, situações nas quais a decisão avaliativa impacta significativamente a vida do sujeito e sejam mais difíceis de serem revertidas, como seleções de emprego, demandam o uso de testes cujos escores sejam bastante precisos. Contudo, para avaliações nas quais o teste específico desempenha um papel não central na avaliação como um todo, a precisão dos escores é menos relevante. Desta maneira, não há justificativas técnicas para coibir o uso, em todas as circunstâncias, de instrumentos cujos escores tenham precisão menor do que 0,80 ou 0,70. Assim, em algumas circunstâncias específicas, nas quais os escores dos testes possam ser complementados com outras informações do examinando, é possível utilizar instrumentos mais modestos em termos de precisão. Nesse sentido, a responsabilidade técnica é do profissional de Psicologia, que deve avaliar qual é o instrumento mais adequado para a demanda específica, ponderando os indicadores de precisão dos escores e o grau de importância desses escores na avaliação psicológica como um todo.

II – Evidências de validade

Segundo o *Standards for Educational and Psychological Testing*, a obtenção de evidências de validade dos testes é o parâmetro mais importante e fundamental no desenvolvimento e avaliação do teste (AERA et al., 2014; Cohen, Swerdlik, & Sturman, 2014). A validade refere-se ao grau em que as evidências empíricas e a teoria suportam as interpretações dos escores do teste para um propósito específico de uso. Este propósito ou contexto específico pode ser, por exemplo, a avaliação de uma criança com dificulda-

des em leitura ou um processo seletivo para vaga em uma organização de trabalho. Quando os escores são interpretados em mais de um contexto, cada interpretação pretendida deve ser validada. Ainda, é importante destacar que quando um usuário/avaliador propõe uma interpretação ou uso dos escores diferentes daqueles propostos pelo desenvolvedor do teste, a responsabilidade por apresentar novas evidências de validade é do usuário/avaliador (AERA et al., 2014).

O entendimento de validade tem passado por modificações desde a primeira versão do *Standards* em 1986. Nos últimos anos, a importância de obtenção de evidências de validade tem sido cada vez mais reconhecida. No entanto, tal processo não tem sido consensual. Por exemplo, Newton (2016) assinala que o modelo atual de cinco tipos de fontes de evidências é incapaz de acomodar todos os tipos de evidências de validade. O autor propõe a distinção entre dois tipos de evidências de validade, a saber: a macrovalidação e a microvalidação.

Apesar das controvérsias, pode-se afirmar que, atualmente, a perspectiva mais aceita na comunidade científica e considerada na Resolução nº 009/2018 é a apresentada na última versão dos *Standards* de 2014 que identifica cinco fontes de evidências, a saber: (1) evidências baseadas no conteúdo do teste; (2) evidências baseadas no processo de resposta; (3) evidências baseadas na estrutura interna; (4) evidências baseadas na relação com outras variáveis; e (5) evidências baseadas nas consequências da testagem.

O primeiro tipo de validade – evidências baseadas no conteúdo do teste – refere-se à análise do relacionamento entre o conteúdo abordado pelos itens do teste e o construto que se pretende medir; questiona-se o quanto os itens constituem amostras do domínio que se pretende avaliar. O segundo tipo – evidências baseadas no processo de resposta – refere-se ao estudo de como os examinandos respondem a um item de um teste; em outras palavras, quais os processos cognitivos necessários e envolvidos para que o examinando responda um determinado item. Esse tipo de análise pode proporcionar evidências em relação a adequação entre o construto e a natureza detalhada do desempenho ou resposta realmente utilizada pelo avaliando. Em relação ao terceiro tipo, a análise da estrutura interna de um teste pode indicar o grau de adequação da relação entre itens do teste e o seu fator (variável-latente). A teoria que embasou a construção de um teste pode sugerir uma única dimensão ou várias dimen-

sões. A técnica da análise fatorial, tanto exploratória quanto confirmatória, tem sido comumente empregada para obtenção de evidências empíricas desse tipo de validade. O quarto tipo – evidências baseadas na relação com outras variáveis – refere-se a evidências baseadas nas relações com outras variáveis, que buscam os padrões de correlação entre os escores do teste e outras variáveis que meçam supostamente o mesmo construto ou construtos relacionados (validade convergente) e com variáveis que meçam construtos diferentes (validade discriminante). Nesta categoria temos, ainda, “relacionamento com o critério”, em que um estudo preditivo indica a força do relacionamento entre o escore e um critério que é obtido tempo depois e um estudo concorrente em que os escores do teste e a informação do critério são obtidas ao mesmo tempo. Por fim, o quinto tipo – evidências baseadas nas consequências da testagem – relaciona o teste às consequências de suas utilizações que podem ser tanto desejadas quanto indesejadas (AERA et al., 2014).

Destaca-se que a Resolução do CFP (009/2018), a partir do Anexo 2, especifica que para um teste ser considerado excelente (Nível A+), ele precisa ter apresentado estudo de estrutura interna (quando aplicável) e três ou mais estudos de diferentes fontes de evidência de validade, com amostras amplas/diversificadas, para testes objetivos. Para testes projetivos, deve-se considerar dois ou mais estudos de diferentes fontes de evidência de validade. Para o mínimo esperado – nível B (suficiente) – é exigido presença de estudo de estrutura interna (quando aplicável) e de pelo menos outro estudo de validade (desde que não seja de conteúdo).

Os critérios de validade abarcados na Resolução nº 009/2018 diferem da Resolução nº 002/2003 principalmente sobre a obrigatoriedade do intervalo de tempo entre os estudos de evidências de validade. O prazo de validade foi modificado de 20 anos (Resolução Nº 002/2003) para 15 anos na nova Resolução (009/2018). Além disso, a Resolução nº 009/2018 especifica a obrigatoriedade do estudo da estrutura interna, quando aplicável. Assim, anteriormente, para um teste ter evidências de validade (nível B – suficiente), ele precisava ter pelo menos um estudo de validade (desde que não seja somente de consistência interna e ou de conteúdo). Atualmente, para obter o mesmo nível B, é necessária a realização do estudo de estrutura interna (quando aplicável), além de um outro estudo de validade (desde que não seja de conteúdo).

III – Sistema de correção e interpretação dos escores

O sistema de correção dos escores é um aspecto central na utilização de testes psicológicos. Falhas nesta etapa comprometem seriamente a análise dos escores do examinando. Basicamente, a proposta é transformar os escores brutos em informações mais interpretáveis. Para tanto, a Resolução nº 009/2018 propõe analisar os testes com base em três possíveis sistemas de correção: referenciado à norma, à teoria (conteúdo, critério e outros tipos) e pontos de corte. Obviamente, casos específicos que não se encaixam nessa taxonomia também podem ser avaliados.

O sistema referenciado à norma é o mais utilizado e diz respeito à comparação dos escores brutos de um indivíduo com a população da qual ele foi retirado. Por exemplo, acertar cinco itens de um teste de inteligência composto de dez questões não respalda muita informação sobre as capacidades cognitivas do examinando. Contudo, se soubermos que a população acerta em média três itens, já teremos, ao menos, uma ideia de que o escore desse indivíduo é superior à média. Além disso, a própria média (ou qualquer outro indicador de tendência central) oferece apenas informações parciais: é possível que o grupo de referência seja bastante heterogêneo e, neste caso, variações em torno da média podem não significar muita diferença em relação ao grupo normativo. Por isso, ao comparar um escore bruto de um indivíduo com a média do grupo, é necessário ponderar a variabilidade do grupo normativo. Isso pode ser realizado por meio de cálculos de escores Z ou variações dele, como T e QI. Além da média, é possível utilizar a distribuição dos escores brutos no grupo de referência para a interpretação. Neste caso, utiliza-se os percentis (ou qualquer outro corte na distribuição), que indica a porcentagem da amostra de referência cujos escores são inferiores ao obtido pelo sujeito.

Destaca-se que mais de um subgrupo de referência pode ser necessário para garantir a justiça na interpretação dos escores de um examinando (AERA et al., 2014). Por exemplo, em um teste de inteligência, a comparação dos escores de uma criança de 10 anos com a população geral de pessoas entre 5 e 60 anos pode não ser justa e comprometer a interpretação dos escores normatizados dessa criança. Nesse sentido, os subgrupos de referência devem ser cuidadosamente definidos e descritos pelos autores de testes psicológicos, que, para tanto, devem ponderar os aspectos teó-

ricos do desenvolvimento psicológico, os condicionamentos sociais e ambientais, entre outros (AERA et al., 2014). Assim, busca-se que os examinandos tenham seus escores interpretados e comparados com populações adequadas às suas características (e não com qualquer grupo), maximizando a justiça na avaliação psicológica.

Destaca-se que a simples análise da diferença de média entre os subgrupos pode não ser suficiente para garantir interpretações justas dos escores. As médias são sensíveis a muitos vieses! Nesse sentido, sugere-se aos autores de testes avaliar, por exemplo, o quanto eventuais diferenças entre grupos normativos são atribuídas, de fato, a diferenças de médias ou estão associadas a vieses dos parâmetros dos itens. Vieses na parametrização dos itens comprometem seriamente a interpretabilidade dos escores, mesmo normatizados (Millsap, & Meredith, 2007; Wu, & Zumbo, 2007). Novamente, os grupos de referência devem ser cuidadosamente analisados.

Além disso, o sistema de interpretação referenciado à norma se apoia em amostras minimamente representativas, de tamanho adequado e atualizadas. Poucos participantes, de uma região específica do país, podem não representar adequadamente a população brasileira, no entanto, podem ser adequados para essa região específica. Ademais, normas estabelecidas em estudos realizados há décadas podem incorrer em injustiças na interpretação dos escores, principalmente se tais escores se referem a construtos que sofrem variação com o tempo ou diferentes gerações.

Além do sistema referenciado à norma, os escores dos examinandos podem ser comparados com algum critério teórico (AERA et al., 2014). Essa forma de interpretação é bastante útil para testes de avaliação do desenvolvimento. Para um instrumento que se destina a avaliar os estágios do desenvolvimento segundo Piaget, por exemplo, os escores brutos do indivíduo podem ser comparados com os critérios teóricos de cada estágio, oferecendo suporte para a classificação do examinando. Neste caso, os escores do examinando não necessariamente devem ser comparados com a média da população brasileira (aliás, a média, neste exemplo, ofereceria pouca informação realmente relevante).

O estabelecimento de normas com referência a critérios deve estar embasado em uma teoria robusta. Nesse sentido, sugere-se que o manual

desse tipo de teste contenha uma descrição bastante detalhada dos aspectos teóricos, que sustentam a interpretação dos escores. Além disso, é necessário apresentar evidências empíricas de que os escores do instrumento, de fato, estão associados aos critérios teóricos (por exemplo, pontos de corte diferenciam adequadamente crianças entre os estágios operatório concreto e formal).

O sistema referenciado a critério também é útil para o estabelecimento de mapas do construto psicológico avaliado. Com base nos parâmetros dos itens e das pessoas, é possível inferir para quais tipos de itens o acerto (ou endosso) é mais provável, dado o nível do construto psicológico (Wilson, 2005). Nesse sentido, associa-se o construto psicológico dos examinandos (estimado a partir de uma amostra) com o conteúdo dos itens. Por exemplo, considere os seguintes itens de autorrelato, de resposta “sim ou não”, destinados a avaliar Extroversão: 1. gosto de sair com os amigos; 2. tenho facilidade para me relacionar com pessoas novas; 3. gosto de ser o centro das atenções. A partir de um estudo empírico, é possível estabelecer o nível de Extroversão de pessoas que tipicamente endossam (ou respondem sim) cada um dos itens. Nesse exemplo, supondo que pessoas com nível mediano de Extroversão tenham maior probabilidade de responder afirmativamente os itens 1 e 2, mas não o 3; é possível interpretar que examinandos medianamente Extrovertidos gostam e tem facilidade para se relacionar com outras pessoas, mas não se sentem à vontade quando os “holofotes estão sobre eles”. O leitor deve ter percebido que esse tipo de estudo auxilia bastante a interpretação dos escores!

Finalmente, o sistema interpretação pode estar embasado em pontos de corte (AERA et al., 2014). Na realidade, esse tipo de correção dos escores também é referenciado ao critério, pois os escores brutos são comparados com alguma referência (ou critério) capaz de classificar os examinandos em grupos teoricamente definidos (por exemplo, com transtorno de personalidade e sem o transtorno). A Resolução nº 009/2018 dedica um ponto específico para essa discussão (B8.3 do Anexo da Resolução), dada a sua relevância para o diagnóstico clínico. Nesse sistema, busca-se estabelecer o ponto de corte ideal dos escores para classificar os examinandos em grupos, teoricamente definidos, da maneira mais precisa possível.

No entanto, classificações diagnósticas sempre são suscetíveis a falsos positivos (quando o resultado

do teste é positivo, mas o examinando não pertence ao grupo ou não tem a patologia) e falsos negativos (quando o resultado do teste é negativo, mas o examinando pertence ao grupo ou tem a patologia). Nesse contexto, é possível investigar a acurácia dos pontos de corte por meio da sensibilidade (taxa de verdadeiros positivos) e especificidade (taxa de verdadeiros negativos). Ressalta-se que não existe um ponto de corte ideal para todas as situações específicas de avaliação psicológica. Para triagem, por exemplo, a sensibilidade é realmente importante, contudo, falsos positivos são toleráveis. Assim, neste exemplo, os pontos de corte podem ser mais brandos, mesmo que isso sacrifique um pouco a acurácia como um todo. Nesse contexto, sugere-se que o profissional de Psicologia avalie se a acurácia dos pontos de corte são adequadas a sua demanda específica de avaliação antes de decidir pela utilização ou não do instrumento!

Além do novo espaço dedicado aos pontos de corte, a nova Resolução nº 009/2018 aperfeiçoou os critérios de avaliação da qualidade do sistema de correção e interpretação dos testes submetidos ao Satepsi. Na Resolução anterior (002/2003), para ser considerado no nível B (suficiente para aprovação), o teste deveria apresentar “amostra em número razoável de sujeitos, permitindo uma apreciação da questão de representatividade”. Já a Resolução de 2018, de maneira mais objetiva, estabeleceu um mínimo de 500 participantes (testes psicométricos). Ressalta-se que se trata de um critério mínimo, no entanto, não necessariamente ideal. Instrumentos de excelência (nível A+), devem garantir amostras maiores (mínimo de 1.000 participantes), das cinco regiões do país e controle de variáveis que podem interferir na interpretação dos escores (por exemplo, idade para a normatização dos escores de um teste de inteligência). Destaca-se que o critério de amostras das cinco regiões (para instrumentos nível A+) é uma ponderação da nova Resolução.

Considerações finais

Conclui-se que, seguindo os avanços que a Resolução nº 002/2003 e posteriores Resoluções e notas técnicas proporcionaram à área da Avaliação Psicológica, a Resolução nº 009/2018 apresenta novas contribuições relevantes para a área da testagem e avaliação psicológica. A partir de um contínuo aprimoramento, o Satepsi buscou critérios psicométricos alinhados aos avanços técnicos científicos da área, inclusive, adotando parâmetros mais elevados para alguns dos tipos de evidências psicométricas. Com isso, esperamos que a nova Resolução impulse os desenvolvedores de testes, bem como a comunidade científica a uma busca de aprimoramento dos testes e, conseqüentemente, do processo de avaliação psicológica.

É igualmente importante atentar para a necessidade de uma formação profissional adequada e constante atualização profissional dos(as) psicólogos(as) nos cursos de graduação de Psicologia. Na literatura é possível encontrar várias publicações que discutem a formação acadêmica e profissional para atuação na área de avaliação psicológica (Bardagi, Teixeira, Segabinazi, Schelini, & Nascimento, 2015; Noronha, & Reppoldi, 2010; Nunes et al., 2012). É de conhecimento que para uma atuação efetiva na área da avaliação psicológica é necessário que um amplo conjunto de competências seja desenvolvido no curso de graduação ou especialização complementar. A problemática parece longe de estar resolvida. Em recente publicação, Zaia, Oliveira e Nakano (2018) objetivaram analisar os processos éticos publicados no jornal do CFP. As autoras verificaram que de 57 processos éticos descritos, 34 foram relacionados à área da avaliação psicológica, representando 60% do total dos respectivos números. Esforços múltiplos de vários atores ainda são requeridos para um efetivo desenvolvimento da área da avaliação psicológica no Brasil.

Referências

- American Educational Research Association – AERA, American Psychological Association – APA, and National Council on Measurement in Education – NCME. (2014). *Standards for educational and psychological testing*. Washington, DC: o autor.
- Andrade, J. M., Laros, J. A., & Gouveia, V. V. (2010). O uso da teoria de resposta ao item em avaliações educacionais: diretrizes para pesquisadores. *Avaliação Psicológica*, 9(3), 421-435.
- Andrade, J. M. de, & Sales, H. F. S. (2017). A diferenciação entre avaliação psicológica e testagem psicológica: questões emergentes. In M. R. C. Lins, & J. C. Borsa (Orgs.), *Avaliação Psicológica: Aspectos Teóricos e Práticos* (pp. 9-22). Petrópolis: Editora Vozes.

- Bardagi, M. P., Teixeira, M. A. P., Segabinazi, J. D., Schelini, P. W., & Nascimento, E. (2015). Ensino da avaliação psicológica no Brasil: levantamento com docentes de diferentes regiões. *Avaliação Psicológica*, 14(2), 253-260.
- Bauer, A., Alavarse, O., & Oliveira, R. (2015). Avaliações em larga escala: uma sistematização do debate. *Educação e Pesquisa*, 41(especial), 1367-1384. <https://doi.org/10.1590/S1517-9702201508144607>
- Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2014). *Testagem e avaliação psicológica: introdução a testes e medidas* (8a ed). Porto Alegre, RS: AMGH.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equations models with unobservable variables and measurement error. *Journal of Marketing*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- Hambleton, R. K., & Swaminathan, H. (1991). *Fundamentals of Item Response Theory*. London, UK: Sage.
- Kamata, A., & Bauer, D. J. (2008). A Note on the Relation Between Factor Analytic and Item Response Theory Models. *Structural Equation Modeling*, 15(1), 136-153. <https://doi.org/10.1080/10705510701758406>
- Millsap, R. E., & Meredith, W. (2007). Factorial invariance: Historical perspectives and new problems. Em R. Cudeck & R. C. MacCallum (Eds.). *Factor analysis at 100: Historical development and future directions* (pp. 131-152). Mahwah, NJ: Lawrence Erlbaum.
- Newton, P. E. (2016). Macro- and micro-validation: Beyond the 'five sources' framework for classifying validation evidence and analysis. *Practical Assessment, Research & Evaluation*, 21, 1-13. Available online: <http://pareonline.net/getvn.asp?v=21&n=12> .
- Noronha, A. P. P., & Reppold, C. T. (2010). Considerações sobre a avaliação Psicológica no Brasil. *Psicologia: Ciência e Profissão*, 30(n especial), 192-201. <https://dx.doi.org/10.1590/S1414-98932010000500009>
- Nunes, M. F. O., Muniz, M., Reppold, C. T., Faiad, C., Bueno, J. M. H., & Noronha, A. P. P. (2012). Diretrizes para o ensino de avaliação psicológica. *Avaliação Psicológica*, 11(2), 309-316.
- Nunnally, J. C. (1994). *Psychometric Theory* (3a ed.). New York, NY: McGraw-Hill.
- Pasquali, L. (2013). *Psicometria: Teoria dos testes na psicologia e na educação* (5a ed). Petrópolis, RJ: Vozes.
- Reppold, C. T. (2011). Qualificação da avaliação psicológica: critérios de reconhecimento e validação a partir dos direitos humanos. In Conselho Federal de Psicologia (Org.), *Ano da avaliação psicológica: textos geradores* (1ª Ed., pp. 21-28). Brasília, DF: Conselho Federal de Psicologia.
- Resolução Nº 002, de 23 de março de 2003*. Define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos e revoga a Resolução CFP nº 025/2001. Brasília, DF: Conselho Federal de Psicologia.
- Resolução Nº 009, de 25 de abril de 2018*. Estabelece diretrizes para a realização de Avaliação Psicológica no exercício profissional da psicóloga e do psicólogo, regulamenta o Sistema de Avaliação de Testes Psicológicos - SATEPSI e revoga as Resoluções nº 002/2003, nº 006/2004 e nº 005/2012 e Notas Técnicas nº 01/2017 e 02/2017. Brasília, DF: Conselho Federal de Psicologia.
- Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema*, 26(1), 108-116. <https://doi.org/doi:10.7334/psicothema2013.260>
- Thompson, B. (2002). *Score reliability: Contemporary thinking on reliability issues*. Thousand Oaks, CA: Sage publications.
- Valentini, F., & Damásio, B. F. (2016). Variância Média Extraída e Confiabilidade Composta: Indicadores de Precisão. *Psicologia: Teoria e Pesquisa*, 32(2), 1-7. <http://dx.doi.org/10.1590/0102-3772e322225>
- Valentini, F., & Laros, J. A. (2011). Teoria de Resposta ao Item na Avaliação Psicológica. In R. A. M. Ambiel, I. S. Rabelo, S. V. Pacanaro, G. A. S. Alves, & I. F. A. S. Leme (Eds.), *Avaliação psicológica: guia de consulta para estudantes e profissionais de psicologia* (pp. 81-108). São Paulo, SP: Casa do Psicólogo.
- Wilson, M. (2005). *Constructing measures: An Item Response Modeling approach*. Mahwah, NJ: L. Erlbaum Associates.
- Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation*, 12(3), 1-26.

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

Josemberg Moura de Andrade

Universidade de Brasília, DF, Brasil. Bolsista de Pesquisa do CEBRASPE/UnB.

E-mail: josemberg.andrade@gmail.com

Felipe Valentini

Universidade São Francisco, SP, Brasil. Bolsista de Pesquisa CEBRASPE/UnB.

E-mail: valentini.felipe@gmail.com

Endereço para envio de correspondência:

Universidade de Brasília (UnB), Campus Universitário Darcy Ribeiro, ICC Sul, Instituto de Psicologia - Asa Norte - Brasília, DF – Brasil. CEP: 70910900

Recebido 21/07/2018

Aprovado 09/08/2018

Received 07/21/2018

Approved 08/09/2018

Recibido 21/07/2018

Aceptado 09/08/2018

Como citar: Andrade, J. M., & Valentini, F. (2018). Diretrizes para a Construção de Testes Psicológicos: a Resolução CFP nº 009/2018 em Destaque. *Psicologia: Ciência e Profissão*, 38(n.spe), 28-39. <https://doi.org/10.1590/1982-3703000208890>

How to cite: Andrade, J. M., & Valentini, F. (2018). Guidelines for the Construction of Psychological Tests: Regulation CFP No: 009/2018 in Highlight. *Psicologia: Ciência e Profissão*, 38(n.spe), 28-39. <https://doi.org/10.1590/1982-3703000208890>

Cómo citar: Andrade, J. M., & Valentini, F. (2018). Directrices para la Construcción de Pruebas Psicológicas: la Resolución CFP nº 009/2018 en Destaque. *Psicologia: Ciência e Profissão*, 38(n.spe), 28-39. <https://doi.org/10.1590/1982-3703000208890>