

A Escala CBCL/6-18 no Brasil: Estudos Exploratórios sobre Evidências de Validade e Precisão

The CBCL/6-18 Scale in Brazil: Exploratory Studies on Evidence of Validity and Accuracy

La Escala CBCL/6-18 en Brasil: Estudios Exploratorios sobre Evidencia de Validez y Precisión

Rauni Jandé Roama-Alves

Universidade Federal de Mato Grosso (UFMT)

Rosimeire de Moraes Amorim

Centro Universitário de Várzea Grande (Univag)

Tatiane Lebre Dias

Rosângela Kátia Sanches Mazzorana Ribeiro

Universidade Federal de Mato Grosso (UFMT)

Raiane Mariah Roama Alves

Faculdade de Pinhais (Fapi)

Carolina Rosa Campos

Universidade Federal do Triângulo Mineiro (UFTM)

Resumo

Introdução: O Inventário de Comportamentos da Infância e Adolescência 6-18 (CBCL/6-18) visa avaliar problemas de comportamento e competência social de crianças e adolescentes de 6 a 18 anos, pelo relato dos pais. Devido à escassez de estudos psicométricos no Brasil, o objetivo deste estudo foi de realizá-los. Foram adotadas análises consideradas representativas da busca de evidências de validade baseadas no conteúdo (EVC), nas relações com variáveis externas (EVE) e na estrutura interna e precisão (EVI). **Método:** participaram das EVC três juízas especialistas, e os dados foram analisados por meio do *Kappa de Fleiss*. Para EVE, a amostra foi composta de 94 pais de crianças e adolescentes, sendo uma parcela sem queixas comportamentais e outra com deficiência intelectual (GDI). Diferenças entre as respostas dos dois grupos foram testadas por meio do teste t. Para as EVI houve uma amostra de 202 pais provenientes de diferentes tipos de escola, à qual se administrou a Análise Fatorial Confirmatória. **Resultados:** foram encontrados resultados satisfatórios às EVE, com diferenças entre os grupos, mas não tão satisfatórios às EVC e EVI. **Discussão:** nas EVE, os grupos se diferenciaram em razão da própria descrição diagnóstica do GDI. As EVC identificaram itens que apresentaram escrita genérica e com baixa qualidade na escrita. Nas EVI houve principalmente índices de ajuste não aceitáveis ao modelo, com indicativos de análises fatoriais exploratórias no país. **Conclusão:** tais resultados apoiam certas evidências de validade do instrumental, mas também reforçam a continuidade de suas investigações.

Palavras-chave: psicopatologia da criança, psicometria, avaliação psicológica

Abstract

Introduction: The Childhood and Adolescent Behavior Inventory 6-18 (CBCL/6-18) aims to assess behavioral problems and social competence of children and adolescents aged 6 to 18 years, as assessed by their parents. Due to the scarcity of psychometric studies in Brazil, the objective of this study was to conduct them. Analyses considered representative of the search for evidence of validity based on content (EVC), relationships with external variables (EVE) and internal structure and precision (EVI) were adopted. **Method:** three expert judges participated in the EVC, and the data were analyzed using Fleiss' Kappa. For EVE, the sample consisted of 94 parents of children and adolescents, one group without behavioral complaints and the other with intellectual disabilities (GDI). Differences between the responses of the two groups were tested using the t-test. For EVI, there was a sample of 202 parents from different types of schools, to whom Confirmatory Factor Analysis was administered. **Results:** satisfactory results were found for the EVE, with differences between the groups, but not so satisfactory for the EVC and EVI. **Discussion:** in the EVE, the groups differed due to the diagnostic description of the GDI itself. The EVC identified items that presented generic writing and low quality in writing. In the EVI, there were mainly unacceptable adjustment indices for the model, with indications of exploratory factor analyses in the country. **Conclusion:** these results support certain evidence of validity of the instrument, but also reinforce the continuity of its investigations.

Keywords: child psychopathology, psychometrics, psychological assessment

Resumen

Introducción: El Inventario de Conducta de la Niñez y la Adolescencia 6-18 (CBCL/6-18) tiene como objetivo evaluar los problemas de conducta y competencia social de niños y adolescentes de 6 a 18 años, por parte de sus padres. Debido a la escasez de estudios psicométricos en Brasil, el objetivo de este estudio fue realizarlos. Se realizaron análisis considerados representativos de la búsqueda de evidencia de validez basada en contenido (EVC), relaciones con variables externas (EVE), estructura interna y precisión (EVI). Método: en la EVC participaron tres jueces expertos. Para el EVE, la muestra estuvo compuesta por 94 padres de niños y adolescentes, algunos sin quejas de conducta y otros con discapacidad intelectual (DIG). Para EVI hubo 202 padres de diferentes tipos de escuelas. Resultados: se encontraron resultados satisfactorios para EVE, con diferencias entre grupos, pero no tan satisfactorios para EVC y EVI. Discusión: en el EVE los grupos se diferenciaron por la propia descripción diagnóstica del GDI. La EVC identificó elementos con redacción genérica y mala calidad de redacción. Para el EVI, hubo principalmente índices de ajuste del modelo inaceptables, lo que indica análisis exploratorios en el país. Conclusión: tales resultados apoyan ciertas evidencias de la validez del instrumento, pero también refuerzan la continuidad de sus investigaciones.

Palabras clave: psicopatología infantil, psicometría, evaluación psicológica.

Introdução

O *Child Behavior Checklist* (CBCL) teve sua primeira versão desenvolvida no ano de 1981, nos Estados Unidos, pelo psiquiatra Thomas Achenbach. Propõe-se avaliar a competência social e problemas de comportamento em crianças e adolescentes por meio do relato dos pais (Achenbach, 1991). Segundo Achenbach (2001), o CBCL é um dos instrumentos mais utilizados em multiculturas, tendo sido traduzido para 55 idiomas. Do mesmo modo, é um dos mais utilizados no Brasil em contextos clínicos e de pesquisa (Borsa & Bandeira, 2011). Sua construção foi empiricamente baseada, não possuindo uma teoria *a priori*, e com uma estrutura interna determinada por análises fatoriais de uma lista de queixas comportamentais, presentes em prontuários médicos.

Posteriormente, após extensivos estudos psicométricos, foram geradas diversas versões, com diferentes itens e para diversas faixas etárias. Atualmente, está disponível em dois formatos: o CBCL para crianças pré-escolares de 1 ano e meio a 5 anos (CBCL 1½-5) e o CBCL para escolares de 6 anos a 18 anos (CBCL/6-18) (Achenbach, 2001; Bordin et al. 2010). Para esse último, alvo deste estudo, sua denominação foi traduzida especificamente como “Inventário de Comportamentos da Infância e Adolescência 6-18” (Duarte & Bordin, 2000) (Achenbach & Rescorla, 2001). Sua última edição foi publicada originalmente em 2001, com a última tradução para o português brasileiro realizada em 2010.

O CBCL/6-18 é composto por 140 itens, que se subdividem em dois domínios, sendo Competência Social, com 20 itens, e Problemas de Comportamento, que englobam 120 itens (Achenbach, 1991; Bandeira et al., 2010, Bordin et al., 1995; Silveiras et al., 2006). O primeiro domínio possui três escalas relativas a problemas no desempenho de atividades, na sociabilidade e na escolaridade. Os itens questionam o envolvimento da criança em brincadeiras, jogos, relações interpessoais, independência no brincar e desempenho escolar (Bordin et al., 1995). Para respondê-los, os pais/responsáveis devem comparar o comportamento do filho com o de outras crianças da mesma idade, considerando as categorias: menos/pior/menor; igual; mais/melhor/maior; não sei (Achenbach, 1991).

O domínio Problemas de Comportamento possui 120 itens que se distribuem em oito escalas sindrômicas, sendo: Ansiedade/Depressão; Isolamento/Depressão; Queixas Somáticas; Problemas Sociais; Problemas de Pensamento; Problemas de Atenção (Desatenção e

Hiperatividade); Comportamento de Quebrar Regras/Delinquencial e Comportamento Agressivo. Além disso, desse total de itens, 17 não se agruparam em nenhum desses oito fatores, sendo incluídos em uma escala denominada de Outros Problemas (Achenbach, 1991; Massola & Silveiras, 2005; Silveiras et al., 2006). As oito escalas sindrômicas, por sua vez, podem ser organizadas em três outros fatores de segunda ordem, hierarquicamente superiores, de acordo com análises fatoriais anteriores, sendo elas: Escala de Problemas de Comportamento Internalizante; Escala de Problemas de Comportamento Externalizante; e Escala Total de Problemas de Comportamento (Achenbach & Rescorla, 2001).

A Escala de Problemas de Comportamento Internalizante corresponde às três primeiras escalas de Problemas de Comportamento (Ansiedade/Depressão, Isolamento/Depressão e Queixas Somáticas), que têm como foco avaliar padrões comportamentais desajustados e problemas emocionais. A Escala de Problemas de Comportamento Externalizante é composta pelas escalas de Comportamento de Quebrar Regras/Delinquencial e Comportamento Agressivo, que possuem itens relacionados a comportamentos inadequados e agressividade (Borsa & Nunes, 2008; Gauy & Guimarães, 2006; Massola & Silveiras, 2005). Por fim, a Escala Total de Problemas de Comportamento abarca todas as oito escalas, mais a escala de Outros Problemas.

Para a avaliação do instrumento, os escores brutos são transformados em *escores T*, que indicam se os indivíduos apresentam comportamentos desviantes de acordo com as normas para idade e sexo, podendo ser classificados como clínicos, limítrofes ou não clínicos. A categoria clínica corresponde a escores baixos, enquanto o oposto se aplica a categoria não clínica. A categoria limítrofe abrange uma faixa intermediária de *escores T*, que indica a necessidade de acompanhamento da criança ou adolescente para identificar um possível aumento nos sintomas e/ou diminuição da competência ao longo do tempo (Bordin et al., 2013).

Interessantemente, verifica-se que os estudos empíricos sobre evidências de validade e precisão da CBCL/6-18 no Brasil são esporádicos e com versões de tradução diferenciadas. Em pesquisa realizada por Roama-Alves e Amorim (2023), buscou-se identificar tais estudos, sendo encontradas apenas quatro produções. O artigo mais antigo foi produzido por Bordin et al. (1995), e possui os seguintes objetivos: (a) avaliar a capacidade discriminativa para critérios diagnósticos da 10ª edição da Classificação Internacional de Doenças CID-10 (1993) e os níveis de gravidade dos transtornos psiquiátricos; (b) avaliar a sensibilidade e especificidade segundo diferentes critérios de teste-positividade; e (c) apontar as modificações introduzidas no questionário durante o processo de tradução e adaptação cultural. Os dados indicaram boa sensibilidade do instrumento não somente entre os casos moderados e graves, mas também entre os casos leves. A edição original utilizada neste estudo foi a norte-americana, de 1991.

Rocha et al. (2013) realizaram estudo que investigou a adequação do modelo de oito escalas sindrômicas do CBCL/6-18 para a população brasileira. As autoras relataram que foi utilizada a edição traduzida de 2001. Os resultados das classificações médias dos itens e das consistências internas da escala foram muito semelhantes aos encontrados nos Estados Unidos e no Uruguai. A Análise Fatorial Confirmatória (AFC) indicou que os dados brasileiros apresentaram o melhor ajuste ao modelo das oito síndromes do que todos os países estudados até aquele momento. Os padrões de gênero foram comparáveis aos relatados em

outros países, mas as pontuações médias de problemas para crianças brasileiras não clínicas foram mais altas do que aquelas para crianças estadunidenses. Portanto, o CBCL discriminou menos entre crianças clínicas e não clínicas no Brasil do que nos Estados Unidos. Dentre as hipóteses explicativas a esse resultado que foram aventadas pelos autores, tem-se que: provavelmente os pais brasileiros podem ter tido uma maior tendência a referenciar problemas de comportamento; variáveis socioeconômicas mais deficitárias poderiam explicar maiores dificuldades comportamentais, como apontado em estudos anteriores do CBCL, padrão esse que pode ter se repetido no Brasil em relação aos Estados Unidos; crianças com queixas de comportamento podem ter sido incluídas na amostra, pois não houve controle tão acurado dessa variável.

Bordin et al. (2013) realizaram uma revisão do CBCL/6-18, do *Youth Self-Report* (YSR) e do *Teacher's Report Form* (TRF), abordando uma visão geral sobre o desenvolvimento das versões originais e brasileira. Descreveram as principais alterações em itens, escalas e pontos de corte na pontuação, nas versões originais de 1991 a 2001, no processo de tradução, retrotradução e adaptação cultural para o Brasil. Todas as decisões finais sobre a tradução de itens para o português foram baseadas em consenso entre os profissionais envolvidos no desenvolvimento das atuais versões brasileiras oficiais dos instrumentos.

Frizzo et al. (2015) avaliaram a estabilidade temporal do CBCL/6-18 e do TRF por meio do teste-reteste. Os autores não identificaram a edição utilizada. Os resultados demonstraram alta confiabilidade, tanto para o CBCL quanto para o TRF (0,87–0,91 e 0,62–0,80, respectivamente, para as escalas Internalizantes, Externalizantes e Total). Esses resultados sugeriram que ambos os instrumentos permaneceram estáveis durante o período de um ano, corroborando sua precisão como em estudos internacionais.

Assim, a fim de contribuir ainda mais para a investigação das evidências de validade e precisão da CBCL/6-18, a presente pesquisa foi proposta. Seu objetivo foi administrar análises psicométricas ao CBCL/6-18 em uma amostra brasileira, que são consideradas representativas da busca de evidências de validade baseadas no conteúdo, nas relações com variáveis externas, na estrutura interna e precisão. Tais análises foram compostas por: estudo de concordância entre juízes; comparação entre grupos (pais/responsáveis respondentes de crianças e adolescentes com desenvolvimento típico e pais/responsáveis respondentes de crianças e adolescentes com deficiência intelectual [DI]); AFC e consistência interna.

Cabe ressaltar que o grupo com DI foi selecionado para esse estudo em razão de seus próprios critérios diagnósticos, pois os sujeitos que a possuem devem apresentar elevadas dificuldades comportamentais (*American Psychiatric Association* [APA], 2014). Há deficits no raciocínio, solução de problemas, planejamento, pensamento abstrato, juízo, aprendizagem acadêmica e aprendizagem pela experiência. Esses prejuízos proporcionam um mau funcionamento adaptativo, de modo que haja maiores dificuldades em atingir padrões de independência pessoal e responsabilidade social em um ou mais aspectos da vida diária, incluindo comunicação, participação social, funcionamento acadêmico ou profissional e independência pessoal em casa ou na comunidade.

Método

A apresentação do método foi dividida em dois subtópicos: (a) Estudo de validade baseado no conteúdo e nas relações com variáveis externas; e (b) Estudo de validade baseado

na estrutura interna e de precisão. Tal decisão foi tomada em razão das análises adotadas para cada estudo, principalmente da AFC, que possui uma maior complexidade estatística, composta, por exemplo, da verificação de uma série de índices de ajuste.

Ressalta-se que o procedimento de coleta de dados será apresentado inicialmente de forma única e abrange os dois estudos. Então, na sequência, serão apresentadas as descrições das especificidades metodológicas de cada um deles.

Procedimento de Coleta de Dados

O projeto foi avaliado pelo Comitê de Ética em Pesquisa da UFMT e apresentou Certificação de Análise e Apreciação Ética (CAAE), sob número suprimido. Para o estudo de validade de conteúdo, as participantes foram selecionadas por meio da análise do Currículo Lattes e convidadas a integrarem a pesquisa via e-mail. As que aceitaram, o consentimento para participação no estudo foi feito por meio do Termo de Consentimento Livre e Esclarecido (TCLE). Em seguida, o “Protocolo para análise dos juízes” foi enviado.

Para o estudo de validade baseado em variáveis externas, foram contactadas uma Escola Estadual do Estado de Mato Grosso e a Associação de Pais e Amigos dos Excepcionais (Apae), de Cuiabá, Mato Grosso. Como procedimentos para os grupos dessas duas instituições, foram enviados em envelope, via crianças ou adolescentes, para que os pais ou responsáveis pudessem responder em casa os seguintes documentos: o TCLE para pais/responsáveis, o CBCL/6-18 e o questionário sociodemográfico. Os que consentiram a participação, após responderem a todos esses documentos, os devolveram à instituição de ensino e foram posteriormente recolhidos pelos pesquisadores.

Estudo de Validade Baseada no Conteúdo e nas Relações com Variáveis Externas

Participantes

Participaram do estudo de validade de conteúdo três juízas especialistas, psicólogas, com doutorado e, no mínimo, dois anos de experiência de atuação na área de Desenvolvimento Humano e Psicopatologia Infantil. Os critérios de inclusão foram: ser psicóloga(o) ou psiquiatra; ter, no mínimo, doutorado e dois anos de experiências na área de Desenvolvimento Humano e Psicopatologia Infantil, de acordo com análise de Currículo Lattes. O critério de exclusão foi de ter relatado a não compreensão de alguma instrução do “Protocolo para análise dos juízes”. Contudo, nenhuma foi excluída.

Para o estudo de busca de evidências de validades baseadas nas relações com variáveis externas, a amostra foi composta por 94 pais de crianças e adolescentes com idades entre seis e 18 anos, dividida em dois grupos: (a) pais de crianças e adolescentes sem queixas de comportamentos (GSQ); e (b) pais de crianças e adolescentes com deficiência intelectual (GDI).

Foram estabelecidos critérios de inclusão e exclusão somente para os pais/responsáveis, focos dessa análise, e não para as crianças e os adolescentes. Assim, os critérios de inclusão para o CSQ foram: ter escolaridade de Ensino Fundamental em diante; possuir filho na faixa etária entre 6 e 18 anos, de acordo com seu próprio relato; possuir filho sem queixas de comportamentos e sem deficiências físicas, de acordo com seu próprio relato; e possuir filho que não fizesse uso de medicamentos psicotrópicos, de acordo com seu próprio relato.

Para o GDI, foi somente modificado o segundo critério para: possuir filho(s) na faixa etária entre 6 e 18 anos com deficiência intelectual (DI), de acordo com seu próprio relato. E como critério de exclusão para ambos os grupos: não compreensão em algum item da CBCL/6-18. No entanto, nenhum participante foi excluído.

O GSQ foi composto especificamente de 72 participantes, sendo mães (n = 60; 83,3%) e pais (n = 12; 16,7%). O GDI apresentou 22 participantes, sendo mães (n = 21; 95,5%) e pai (n = 1; 4,5%). A análise de comparação não indicou diferenças entre os grupos quanto a esta variável de gênero dos respondentes ($\chi^2 = 2,078$; p = 0,183; df = 1; V = 0,149, pequeno tamanho de efeito). Também não houve diferença na escolaridade desses respondentes, sendo a maioria composta pela formação em Ensino Superior (GSQ: n = 34; 47,2%; GDI: n = 10; 45,5%) em ambos os grupos ($\chi^2 = 0,05$; p = 0,97; df = 2; V = 0,02, pequeno tamanho de efeito).

Foram também averiguadas as características amostrais das crianças e adolescentes aos quais os pais referiram as respostas no CBCL/6-18. O GSQ foi composto da seguinte forma: meninas: n = 32 (44,4%); meninos: n = 40 (55,6%); idade: mínima = 9 anos, máxima = 17 anos, média = 14,66, DP = 1,81. O GDI apresentou: meninas: n = 5 (22,7%); meninos: n = 17 (77,3%); idade: mínima = 10 anos; máxima = 18 anos; média = 15,09, DP = 2,36. As análises de comparação não indicaram diferenças entre os grupos quanto ao gênero ($\chi^2 = 3,330$; p = 0,084; df = 1; V = 0,188, pequeno tamanho de efeito) e à idade (teste t = -0,891; df = 92; p = 0,375; d = 0,22, pequeno tamanho de efeito).

Instrumentos

CBCL/6-18. Versão traduzida da original de 2001 por Bordin et al. (2010). O inventário é composto por 140 itens, sendo 20 que avaliam Competência Social e 120 que avaliam Problemas de Comportamentos. Deve ser respondido por pais ou responsáveis. Subdivide-se em oito escalas sindrômicas, construídas a partir de análises fatoriais: Ansiedade/Depressão (13 itens); Isolamento/Depressão (8 itens); Queixas Somáticas (11 itens); Problemas Sociais (11 itens); Problemas de Pensamento (15 itens); Problemas de Atenção (Desatenção e Hiperatividade) (10 itens); Comportamento de Quebrar Regras/Delinquencial (17 itens); Comportamento Agressivo (18 itens). O inventário ainda possui 17 itens (Outro Problemas) que não se agruparam em nenhuma desses oito fatores. A partir dos escores obtidos nessas escalas, a criança ou adolescente pode ser incluído nas faixas clínica, limítrofe ou não clínica, em relação ao seu funcionamento global e perfis Internalizante e Externalizante. Tais perfis foram também construídos a partir de análises fatoriais hierárquicas, que os incluíram em segunda ordem, hierarquicamente superiores. Para inclusão no perfil Internalizante são considerados os fatores Ansiedade/Depressão, Isolamento/Depressão e Queixas Somáticas. Para inclusão no perfil Externalizante, os fatores incluídos são Comportamento de Quebrar Regras/Delinquencial e Comportamento Agressivo (Bordin et al., 1995).

Protocolo para análise dos juízes. O CBCL/6-18 possui escalas orientadas pelo Manual Diagnóstico e Estatístico dos Transtornos Mentais (DSM). É composto por oito escalas sindrômicas e a escala Outros Problemas, totalizando nove elementos fundamentais ou escalas fatoriais. Descreveu-se a definição teórica de cada um desses elementos, conforme o DSM-5 (APA, 2014). Foram apresentados todos os itens do inventário (1 ao 120), e ao lado de cada um havia um espaço em branco para que o juiz incluísse o elemento fundamental que julgas-

se representá-lo, de acordo com as definições. Foi também realizada avaliação da adequação dos itens pelos juízes, na qual se apresentou novamente todos os itens e, ao lado, duas colunas: a primeira com o título de “Adequado”, com as opções de 0, 1 e 2, e a segunda com o título “Motivo”, com as opções de 0, 1 e 2. Para seu preenchimento, foi, então, solicitado ao juiz que indicassem se o item é adequado ou não para a avaliação de comportamento infantil, numa escala de 0 a 2, na qual: 0 = não, 1 = parcialmente, 2 = sim. Se marcassem 0 ou 1 em algum item, solicitava-se indicar o motivo, podendo ser mais de um: 1 = a escrita não está clara, 2 = o item é inadequado para as idades do teste, 3 = o item está repetido.

Questionário Sociodemográfico. O instrumento foi elaborado para a presente pesquisa. Foi composto por perguntas fechadas com o objetivo de obter informações sobre características dos participantes, tais como: escolaridade e gênero dos pais; escolaridade, idade e gênero das crianças e adolescentes.

Análise de Dados

Para o estudo de validade de conteúdo, foram realizadas análises estatísticas descritivas e inferenciais por meio dos softwares *Statistical Package for Social Sciences 2.1** (IBM Corp. Released, 2012) e *Excel** (Microsoft Corporation, 2018). Os dados foram submetidos à análise de concordância por meio da medida *Kappa* de Fleiss (k). Para análise de adequabilidade de k, foi seguida a classificação de Landis e Koch (1977): < 0 = Sem Concordância; entre 0 e 0,20 = Concordância Fraca; entre 0,21 e 0,40 = Concordância Justa; entre 0,41 e 0,60 = Concordância Moderada; entre 0,61 e 0,80 = Concordância Substancial; entre 0,81 e 1,00 = Concordância Quase Perfeita.

Para o estudo de validade baseado em variáveis externas, foram utilizados os mesmos softwares e realizadas análises de comparações entre o GSQ e o GDI. Foram comparados todos os itens da CBCL/6-18, bem como suas nove escalas fatoriais. Considerando o tamanho amostral, optou-se pelo uso da estatística paramétrica, sendo utilizada a análise de comparação do teste t para os dados numéricos. Para comparação dos dados nominais, foi utilizado o teste de Qui-quadrado (χ^2).

Para análise de tamanho do efeito dos dados numéricos, foi utilizado o “d de Cohen” (d) com os seguintes valores de referência: próximos a 0,2 como pequeno efeito; próximos a 0,5 como médio efeito; e próximos a 0,8 como grande efeito (Cohen, 1988). Para tamanho de efeito dos dados nominais, foi utilizado o “V de Cramer” (V). Foram utilizados os seguintes valores de referência: pequeno efeito: 0,1 (quando 1 grau de liberdade (df)), 0,07 (2 df), 0,06 (3 df); médio efeito: 0,3 (1 df), 0,21 (2 df), 0,17 (3 df); grande efeito: 0,5 (1 df), 0,35 (2 df), 0,29 (3 df) (Cohen, 1988). O valor de significância (p) adotado para todas as análises foi de $p \leq 0,05$.

Estudo de Validade Baseada na Estrutura Interna e de Precisão

Participantes

A amostra do estudo de evidências de validades baseadas nas relações com variáveis externas foi na sequência ampliada para o presente estudo, em razão da AFC demandar um maior número de participantes (Hair et al., 2009). Por fim, foi composta de 202 pais de crianças e adolescentes, sendo: mães: n = 157 (77,7%); pais: n = 45 (22,3%); escolaridade: Ensino

Fundamental incompleto (n = 44, 21,8%), Ensino Fundamental completo (n = 49, 24,3%), Ensino Médio (n = 47, 23,3%) e Ensino Superior (n = 62, 30,7%). Em relação às crianças e aos adolescentes: meninas: n = 89 (44,1%); meninos n = 113 (55,9%); idade: idade: mínima = 9 anos; máxima = 18 anos; média = 12,97, DP = 2,65. Na Tabela 1 é possível verificar a distribuição por ano escolar das crianças e adolescentes da amostra.

Tabela 1

Distribuição do Ano Escolar das Crianças e Adolescentes da Amostra

Ano escolar	n	%
1º Ano Ensino Médio	23	11,4
2º Ano Ensino Médio	31	15,3
4º Ano Ensino Fundamental	28	13,9
5º Ano Ensino Fundamental	31	15,3
6º Ano Ensino Fundamental	26	12,9
7º Ano Ensino Fundamental	6	3,0
8º Ano Ensino Fundamental	4	2,0
9º Ano Ensino Fundamental	31	15,3
Somente Apae	19	9,4
1º Ano Ensino Médio e Apae	1	0,5
2º Ano Ensino Médio e Apae	1	0,5
7º Ano Ensino Fundamental e Apae	1	0,5
Total	202	100,0

Nota. n: frequência, %: porcentagem.

Instrumentos

Foram utilizados dois dos instrumentos descritos anteriormente: o CBCL/6-18 e o Questionário Sociodemográfico.

Análise de Dados

Para o estudo de precisão, foi utilizado o software livre R, em sua interface R Studio (R Development Core Team, 2017). Para estatística descritiva, foram utilizados os pacotes *fields* e *Hmisc*. Foi adotada a técnica de “parcela de itens” na presente AFC, principalmente em decorrência do tamanho amostral, que impossibilitou a construção de modelos, considerando a análises de todos os itens (Coffman & MacCallum, 2005). Nesse sentido, os itens correspondentes a cada uma das nove escalas foram somados, de modo a se obter um total de cada uma delas, de acordo com o apresentado no manual (Achenbach & Rescorla, 2001). Os modelos testados buscaram identificar se determinados grupos de escalas corresponderiam às variáveis latentes de comportamentos “Internalizantes” (Ansiedade/Depressão, Isolamento/Depressão e Queixas Somáticas) e “Externalizantes” (Comportamento de Quebrar Regras/Delinquencial e Comportamento Agressivo), de acordo com os estudos anteriores. Testou-se também um outro modelo para ambos os fatores, que foram somados a um fator composto pelas outras escalas restantes, denominado de “Outras escalas” (escalas de Problemas Sociais, Problemas relacionados à pensamento, Problemas relacionados à atenção e Outros Problemas). Por fim, também foi testado um terceiro modelo, com a mesma estrutura do

anteriormente citado, porém, sem a escala de Outros Problemas.

Para avaliar a qualidade do ajustamento dos modelos construídos, via AFC, utilizaram-se os seguintes índices (Hair et al., 2009): significância de Qui-quadrado (χ^2): $p > 0,05$; o próprio valor de χ^2 , que quanto maior, pior o ajustamento do modelo; Índice de Ajuste Comparativo (Comparative Fit Index, CFI) $> 0,95$; Índice de Tucker-Lewis (TLI) $> 0,95$; Raiz Quadrada da Média do Erro de Aproximação (Root Mean Square Error of Approximation, RMSEA) $< 0,06$; Intervalo de Confiança [IC] em 90%; e Raiz Quadrada Média Residual Padronizada (Standardized Root Mean Square Residual, SRMR) $< 0,08$. Foram, ainda, considerados aceitáveis valores CFI superiores a 0,90, e valores de SRMR e RMSEA até 0,10 (Hu & Bentler, 1998). Foram também utilizados índices que permitem a comparação entre modelos: Akaike Information Criterion (AIC) e Bayesian Information Criterion (BIC). Nesses casos, quanto menor o valor apresentado dentre os modelos, melhor o ajustamento (Byrne, 2016). Os pacotes adotados para todas essas análises foram: *lavaan*, *semPlot*, *clusterGeneration* e *knitr*.

Os parâmetros de interpretação das Correlações de Pearson (r) adotados foram de: 0 a 0,3 como fraca; 0,3 a 0,7 como moderada; e $\geq 0,7$ como forte (Cronk, 2017). O corte das cargas fatoriais foi estipulado em 0,40 (Damásio, 2012). Para consistência interna, foi utilizado o Ômega de McDonald hierárquico (ω_h) para o teste total e para cada fator independentemente foi utilizado o *alfa* de Cronbach (α). Adotou-se o valor de $\geq 0,7$ como indicativo de boa consistência (Cronk, 2017). Para essa última análise, foram utilizados especificamente os pacotes *psych* e *userfriendlyscience*.

Resultados

Para avaliação da representação do elemento fundamental, as análises das três juízas indicaram que 47 itens da CBCL apresentaram concordância quase perfeita ($k = 1$) e 54 foram considerados com concordância justa pelas especialistas ($k = 0,33$). Os demais itens ($n = 19$) não obtiveram concordância entre as juízas, tendo sido alocados em definições diferenciadas. Em relação à adequação, nenhum dos itens foi considerado como inadequado pelas juízas, porém houve concordância de que três deles eram parcialmente adequados ($k = 0,33$). Para esses três, os motivos elencados pela maioria das juízas (66%) era de que não possuíam escrita clara.

Após o estudo de evidências de validade de conteúdo, buscou-se investigar a validade baseada nas relações com variáveis externas por meio da comparação entre os grupos GSQ e o GDI. Observou-se que 60 itens apresentaram diferença com magnitude de efeito $d \geq 0,50$. Esses itens eram pertencentes às oito escalas sindrômicas, bem como à escala Outros Problemas. Também, com o objetivo de verificar especificamente as diferenças no total dessas escalas, foram comparadas as médias obtidas em cada uma delas entre ambos os grupos, conforme apresentado na Tabela 2 (estatística descritiva obtida e os resultados das comparações).

Tabela 2

Média, Desvio Padrão e Comparação Entre os Grupos GSQ e GDI nas Escalas da CBCL-6-18

Escalas	GSQ		GDI		t	p	d
	M	DP	M	DP			
AnD	5,09	3,39	4,59	2,06	0,66	0,51	-0,16
AfD	1,37	1,11	4,27	2,47	-7,74	≤ 0,001*	2,03
QS	3,65	4,03	3,27	2,47	0,41	0,67	-0,01
PS	1,97	1,94	6,05	1,94	-8,58	≤ 0,001*	2,30
PRP	1,58	2,36	3,86	3,99	-3,32	0,001*	0,91
PRA	4,23	4,12	7,96	2,55	-3,98	≤ 0,001*	1,07
CQR	1,08	1,72	2,05	1,09	-2,46	0,016*	0,65
CA	3,29	2,49	5,68	3,76	-3,46	0,001*	0,91
OP	3,11	2,46	7,14	4,67	-5,31	≤ 0,001*	1,39

Nota: GSQ: Grupo de Criança Sem Queixas de Comportamento; GDI: Grupo de Criança com Deficiência Intelectual; M: Média; DP: Desvio Padrão; t: Teste t de Student; p: Valor de Significância; d: Cohen d; AnD: Ansiedade e Depressão; AfD: Afastamento e Depressão; QS: Queixas Somáticas; OS: Problemas Sociais; PRP: Problemas Relacionados à Pensamento; PRA: Problemas Relacionados à Atenção; CQR: Comportamento de Quebrar Regras; CA: Comportamento Agressivo; OP: Outros Problemas.

De acordo com a Tabela 2, nota-se que somente as escalas de Ansiedade e Depressão e de Queixas Somáticas não apresentaram diferenças significativas entre os grupos. Todas as outras apresentaram diferença com valores de “p” menores que 0,05 e magnitude de efeito $d \geq 0,50$.

Dando continuidade aos objetivos, buscou-se verificar a estrutura interna do instrumento por meio da AFC. Para isso, inicialmente, foram obtidos dados da estatística descritiva de média, desvio padrão, mínimo e máximo das pontuações obtidas nas escalas do CBCL/6-18, que podem ser observados na Tabela 3. As análises de Kolmogorov-Smirnov ($D = 0,10$, $p = 0,99$) e Shapiro-Wilk ($W = 0,97$, $p = 0,91$) indicaram, conjuntamente, normalidade para os dados ($p > 0,05$). Desse modo, o método adotado para as análises foi o de Máxima Verossimilhança (ML), indicado para distribuições como essa (Byrne, 2016).

Tabela 3

Média, Desvio Padrão, Mínimo e Máximo das Pontuações Obtidas nas Escalas do CBCL/6-18

Escalas	Mínimo	Máximo	M	DP
Ansiedade e depressão	0	12	4,43	2,92
Afastamento e depressão	0	10	1,75	2,16
Queixas somáticas	0	10	2,77	3,40
Problemas sociais	0	8	2,12	2,25
Problemas relacionados a pensamento	0	11	1,53	2,48
Problemas relacionados à atenção	0	13	3,50	3,55
Comportamentos de quebrar as regras	0	9	1,07	1,41
Comportamento agressivo	0	21	3,93	3,43
Outros problemas	0	17	3,69	3,28

Nota. M: Média; DP: Desvio Padrão.

Em seguida, testou-se o ajustamento de três modelos à amostra total, a partir dos fatores de segunda ordem, hierarquicamente superiores, sendo o primeiro organizado por: (a) Internalizantes (escalas de Ansiedade/Depressão, Isolamento/Depressão e Queixas Somáticas); e (b) Externalizantes (escalas de Comportamento de Quebrar Regras/Delinquencial e Comportamento Agressivo). O segundo, considerando os mesmos fatores deste primeiro modelo somado a um fator composto pelas outras escalas restantes, denominado de “Outras Escalas” (escalas de Problemas Sociais, Problemas de Pensamento, Problemas de Atenção e Outros Problemas). Um terceiro modelo foi testado com essa mesma estrutura do segundo, porém, sem a escala de Outros Problemas. Os índices de ajuste para cada modelo podem ser observados na Tabela 4.

Tabela 4

Índices de Ajuste dos Modelos de AFC Investigados

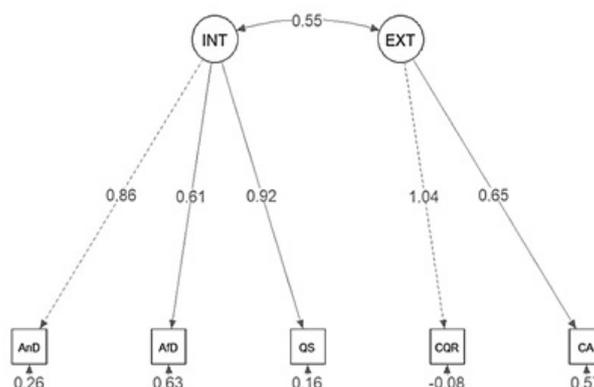
Modelos										
Índices de ajuste para cada modelo										
	RMSEA	IC90%	CFI	TLI	χ^2	p	df	SRMR	AIC	BIC
1	0,371	0,315 – 0,431	0,806	0,514	115,383	≤ 0,001	4	0,122	4294,871	4331,262
2	0,341	0,318 – 0,365	0,684	0,526	588,047	≤ 0,001	24	0,079	7527,431	7596,905
3	0,360	0,332 – 0,388	0,692	0,493	460,834	≤ 0,001	17	0,108	6694,101	6756,958

Nota. 1. Internalizantes, Externalizantes; 2. Internalizantes, Externalizantes e Outras Escalas; 3. Internalizantes, Externalizantes e Outras Escalas (sem Outros Problemas); RMSEA: *Root Mean Square Error of Approximation*; IC90%: Intervalo de Confiança; CFI: *Comparative Fit Index*; TLI: Índice de Tucker-Lewis; χ^2 : Qui-quadrado; p: significância; df: graus de liberdade; SRMR: *Standardized Root Mean Square Residual*; AIC: *Akaike Information Criterion*; BIC: *Bayesian Information Criterion*

Assim, verificou-se índices de ajuste não aceitáveis aos três modelos. Embora o primeiro modelo tenha apresentado os melhores índices, nas análises das cargas fatoriais, no fator “externalizantes”, houve carga maior que um na escala de “comportamentos de quebrar as regras”, indicando um caso de *Heywood* (Kolenikov & Bollen, 2012), o que não é adequado em análises como a realizada. As outras cargas fatoriais, por sua vez, puderam ser consideradas adequadas, como podem ser observadas na Figura 1.

Figura 1

Representação Gráfica do Primeiro Modelo com os Melhores Índices Testados



Em relação à consistência interna, para o CBCL/6-18 total, com todas as escalas, foi identificado valor de $\omega h = 0,74$. Para o fator Internalizantes: $\alpha = 0,82$; para o Externalizantes: $\alpha = 0,65$; para o fator Outras Escalas: $\alpha = 0,88$; para o fator Outras Escalas sem a escala de Outros Problemas: $\alpha = 0,84$.

Discussão

Na construção da escala de perfil do comportamento do CBCL, Achenbach (1991), para elaborar itens, procurou seguir um rigor metodológico que representasse de fato os problemas de comportamentos internalizantes e externalizantes da criança e do adolescente. Desse modo, esse estudo buscou verificar, primeiramente, a validade desses conteúdos no Brasil. Os dados foram submetidos inicialmente à análise de concordância de juízes especialistas por meio da medida *Kappa* de Fleiss. Buscou-se compreender se os itens representariam os domínios que pretendiam avaliar, se estavam adequados em relação à avaliação de comportamento infantil e, caso não houvesse julgados como não adequados, o motivo (baixa qualidade da escrita, não ser adequado para as idades do teste, ser repetido).

O resultado da análise *Kappa* de Fleiss apontou que, em relação aos 120 itens totais que compõem a escala de perfil de comportamento, foram verificadas as concordâncias de “quase perfeita” e de “justa” para a maioria deles ($n = 101$; 84%), quanto à representatividade do elemento fundamental. Como hipótese para os 19 itens que não obtiveram concordância entre as juízas, sugere-se que provavelmente apresentaram uma escrita genérica e não evidente para que fosse possível identificá-los nos elementos. Conforme recomendações feitas por De Vellis (2003) e Pasquali (2010), na construção dos itens, o pesquisador tem que se preocupar com uma variedade de parâmetros que regulem a definição de cada item e da escala como um todo. Além disso, deve assegurar uma redação apropriada, objetiva, clara e a garantia da variabilidade da resposta para um mesmo construto.

Como exemplo de itens não concordantes, pôde-se observar que os itens seis (“Faz cocô na calça ou fora do vaso sanitário”), sete (“Gosta de contar vantagem”) e 24 (“É difícil para comer”) pertencem à uma mesma escala fatorial (Outros Problemas). Por sua vez, os itens 18 (“Machuca-se de propósito ou já tentou se matar”), 70 (“Vê coisas que não existem [descreva]”) e 76 (“Dorme menos que a maioria das crianças ou adolescentes”) também pertencem a uma mesma escala fatorial (Problemas relacionados à pensamento). Apesar de serem escalas separadas, são dois fatores próximos conceitualmente, em que esses itens foram situados em ambos ou até mesmo em outras escalas fatoriais pelas juízas. Os comportamentos/sintomas descritos podem pertencer e serem decorrentes de um mesmo construto e/ou uma condição psiquiátrica. Tal dado leva ao enfraquecimento dos conteúdos teóricos defendidos para os itens e a escala.

Na avaliação dos itens quanto à sua adequação para avaliação do comportamento infantil, 117 deles se mostraram adequados, com concordância classificada como “quase perfeita”. Para os três itens restantes, as especialistas apontaram que a escrita não estava clara. Desse modo, provavelmente esse dado indica uma falha na análise de validade aparente/semântica, que envolve averiguar a opinião do público-alvo e dos especialistas sobre os itens (Urbina, 2014). Segundo Pasquali (1998), na análise semântica, o pesquisador poderá explicar ao público-alvo o que ele pretende dizer com tal item, e os participantes poderão sugerir

como se deveria formular o item para expressar o que o pesquisador quer dizer com ele, possibilitando reformulações.

Como visto, a validade de conteúdo diz respeito a se os itens que compõem o teste são abrangentes e representativos dos construtos que pretendem avaliar. Assim, como outras evidências de validade, é importante que esta também seja assegurada (American Educational Research Association [Aera] et al., 2014). Porém, o presente estudo não demonstrou tal evidência para uma quantidade considerável de itens ($n = 22$). Desse modo, sugere-se que estudos mais robustos sobre essa validade sejam realizados com o instrumento no país, visto que não foi realizado para sua publicação (Roama-Alves & Amorim, 2023).

No presente estudo, buscou-se também verificar a validade baseada nas relações com variáveis externas por meio da análise de comparação entre grupos clínico (com DI) e não clínico (sem queixas de comportamentos e sem deficiências físicas). Assim, os resultados demonstraram que o grupo com DI apresentou resultados inferiores em 60 itens no CBCL. Além disso, no total das nove escalas, somente em duas não houve diferenças entre os grupos, nas de Ansiedade e Depressão e de Queixas Somáticas. Todavia, aventa-se a hipótese de que provavelmente esse resultado tenha sido encontrado, pois crianças com DI podem apresentar frequentemente características internalizantes semelhantes às que tem desenvolvimento típico (Rodríguez, 2004; Rodrigues & Alchieri, 2009). De toda forma, pode-se considerar que esse resultado foi relevante no CBCL, indicando certa capacidade discriminativa do instrumento e a evidência de validade estudada podendo ser reforçada, em concordância com estudo anterior (Bordin et al., 1995).

Por fim, foram buscadas evidências de validade baseada na estrutura interna da CBCL/6-18 por meio da AFC (APA, 2014). Para isso, testou-se o ajustamento de três modelos, com base em estudos anteriores (Achenbach & Rescorla, 2001; Achenbach & Edelbrock, 1983). Como resultado, verificou-se índices de ajuste não aceitáveis para os três. Dentre eles, o com melhores índices foi composto de dois fatores: (a) Internalizantes (composto pelas escalas de Ansiedade e Depressão, Afastamento e Depressão e Queixas somáticas); e (b) Externalizantes (composto pelas escalas de Comportamentos de quebrar as regras e Comportamento Agressivo). Além disso, nas análises das cargas fatoriais, no fator “externalizantes”, houve carga maior que um na escala de “Comportamentos de quebrar as regras”, indicando um caso de Heywood, o que não é adequado em análises como a realizada. Segundo Kolenikov e Bollen (2012), quando ocorre um caso de Heywood, os pesquisadores precisam determinar o motivo de sua ocorrência.

Para o presente caso, levantou-se a hipótese de que houve uma má especificação do modelo, provavelmente em razão dele não aplicável ao nosso contexto. Assim, análises exploratórias podem ainda ser necessárias no Brasil, bem como um melhor alinhamento da validade de conteúdo a essa realidade (Sato, 1987; Bollen, 1989). Porém, cabe ressaltar que os resultados aqui encontrados foram incompatíveis com estudo anteriormente realizado no Brasil, sobre a estrutura fatorial, que indicou muito bons resultados para a AFC em uma amostra de 1228 participantes (Rocha et al., 2013). Contudo, neste estudo, o número de itens apresentado esteve incompatível com a tradução referenciada e a análise não foi realizada para os fatores Internalizantes e Externalizantes, mas para cada uma das oito escalas sindrômicas (Rocha et al., 2013).

Quanto à precisão, os resultados aqui encontrados, por meio da análise de consistência interna, identificaram bons valores ($> 0,7$) para o total da escala e para os fatores Internalizantes

e de Outras Escalas (Cronk, 2017). Fez-se notar que somente para o fator Externalizantes houve valor encontrado abaixo do adotado como indicativo de boa consistência no presente estudo (0,65). Entretanto, autores apontam que valores acima de 0,6 podem ainda ser aceitáveis para boa consistência (Martin & Savage-McGlynn, 2013; Taber, 2018).

Considerações Finais

De modo geral, pode-se observar que embora o CBCL/6-18 seja amplamente utilizado, ainda há poucos estudos que reforcem suas evidências de validade no Brasil. O presente estudo buscou contribuir para uma mudança nesse panorama, e de fato reforçou suas evidências baseadas nas relações com variáveis externas. Para suas evidências de validade baseadas no conteúdo e na estrutura interna, os resultados se mostraram não muito favoráveis.

Todavia, é importante ressaltar as limitações deste estudo. Dentre eles, o mais importante foi o baixo número de participantes na análise de juízes e da AFC. Estudos futuros necessariamente deverão ampliar tais amostras, bem como poderão controlar variáveis como a de gênero para compreensão do modelo fatorial, em uma análise, por exemplo, multigrupo. Os dados demonstraram inclusive que possivelmente os dois grandes fatores do instrumento, Internalizantes e Externalizantes, podem não possuir a mesma estrutura de estudos internacionais, sendo necessária revisão de sua estrutura por meio de uma análise exploratória, caso não seja novamente identificada em uma amostra maior.

Referências

- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. University of Vermont.
- Achenbach, T. M. (2001). *Manual for the Child Behavior Checklist and 1991 profile*. University of Vermont.
- Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. University of Vermont.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. University of Vermont.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. AERA.
- American Psychiatric Association [APA]. (2014). *Manual diagnóstico e estatístico de transtornos mentais: DSM-5*. (5ª ed.). Artmed.
- Bandeira, D. R., Borsa, J. B., Artech, A. X., & Segabinazi, J. D. (2010). Avaliação de problemas de comportamento infantil através do Child Behavior Checklist (CBCL). In C. S. Hutz (Org.), *Avanços em Avaliação Psicológica e Neuropsicológica de crianças e adolescentes* (pp. 101–122). Casa do Psicólogo.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley and Sons, Inc., New York. <https://doi.org/10.1002/9781118619179>
- Bordin, I. A. S., Mari, J. J., & Caeiro, M. F. (1995). Validação da versão brasileira do Child Behavior Checklist (CBCL) (Inventário de Comportamentos da Infância e Adolescência): Dados preliminares. *Revista da Associação Brasileira de Psiquiatria—Asociación*

- Psiquiatria de la America Latina*, 17(2), 55–66. https://www.researchgate.net/publication/285968522_Validation_of_the_Brazilian_version_of_the_Child_Behavior_Checklist_CBCL
- Bordin, I. A. S., Rocha, M. M., Paula, C. S., Teixeira, M. C. T. V., Achenbach, T. M., Rescorla, L. A., & Silveiras, E. F. M. (2013). Child Behavior Checklist (CBCL), Youth Self-Report (YSR) e Tache's Report Form (TRF): Uma visão geral sobre o desenvolvimento das versões originais e brasileiras. *Cadernos de Saúde Pública*, 29(1), 13–28. <https://doi.org/10.1590/S0102-311X201300010004>
- Bordin, I. A. S., Silveiras, E. F. de M., Rocha, M. M., Teixeira, M. C. T. V., & Paula, C. S. (2010). *Inventário de Comportamentos da Infância e adolescência (CBCL 6/18)*. ASEBA.
- Borsa, J. C., & Bandeira, D. R. (2011). Uso de instrumentos psicológicos de avaliação do comportamento agressivo infantil: Análise da produção científica brasileira. *Avaliação Psicológica*, 10(2), 193–203. http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712011000200010&lng=pt&nrm=iso&tlng=pt
- Borsa J. C., & Nunes M. L. T. (2008). Concordância parental sobre problemas de comportamento infantil através do CBCL. *Paidéia*, 18, 317–330. <https://doi.org/10.1590/S0103-863X2008000200009>
- Byrne, B. M. (2016). *Structural equation modeling with Amos: Basic concepts, applications, and programming* (3ª ed.). Routledge.
- Coffman, D. L., & MacCallum, R. C. (2005). *Using parceling to increase statistical power in structural equation modeling*. In T. A. Little, J. A. Schnabel, & J. A. Baumert (Eds.), *Modeling longitudinal and multiple group data: Practical issues, applications, and recommendations* (pp. 235–255). Lawrence Erlbaum Associates.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2ª ed.). Lawrence Erlbaum Associates.
- Cronk, B. C. (2017). *How to use SPSS: A step-by-step guide to analysis and interpretation* (8ª ed.). Routledge.
- Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, 11(2), 213–228. http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712012000200007&lng=pt&tlng=pt
- De Vellis, R. F. (2003). *Scale Development: Theory and Applications* (2ª ed., Vol. 26). Sage Publications.
- Duarte, C. S., & Bordin, I. A. (2000). Instrumentos de avaliação. *Brazilian Journal of Psychiatry*, 22, 55–58. <https://doi.org/10.1590/S1516-44462000000600015>
- Frizzo, G. B., Pedrini, J. R. de Souza, D. S., Bandeira, D. R., & Borsa, J. C. (2015). Reliability of child behavior checklist and teacher's report form in a sample of brazilian children. *Universitas Psychologica*, 14(1), 149–156. <https://doi.org/10.11144/Javeriana.upsy14-1.rcbc>
- Gauy, F. V., & Guimarães, S. S. (2006). Triagem em saúde mental infantil. *Psicologia: Teoria e Pesquisa*, 22, 5–16. <https://doi.org/10.1590/S0102-37722006000100002>
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. E. (2009). *Análise Multivariada de Dados* (6ª ed., 147 p). Bookman.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–

453. <https://doi.org/10.1037/1082-989X.3.4.424>
- IBM Corp. Released (2012). *IBM SPSS Statistics for Windows, Version 21.0*. IBM Corp.
- Kolenikov, S., & Bollen, K. A. (2012). Testing negative error variances: Is a Heywood case a symptom of misspecification? *Sociological Methods & Research*, 41(1), 124–167. <https://doi.org/10.1177/0049124112442138>
- Martin, C., & Savage-McGlynn, E. (2013). A “good practice” guide for the reporting of design and analysis for psychometric evaluation. *Journal of Reproductive and Infant Psychology*, 31(5), 449–455. <https://doi.org/10.1080/02646838.2013.835036>
- Massola, M. C., & Silveiras, E. B. (2005). Análise das escalas sindrômicas do CBCL/6-18: Uma revisão crítica. *Psicologia: Teoria e Pesquisa*, 21(4), 297–309. <https://doi.org/10.1590/S0102-37722005000400005>
- Microsoft Corporation. (2018). *Microsoft Excel*. <https://office.microsoft.com/excel>
- Pasquali, L. (1998). *Psicometria: Teoria e aplicações*. Vozes.
- Pasquali, L. (2010). *Psicometria: Teoria e aplicações* (2ª ed.). Vozes.
- R Development Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org>
- Roama-Alves, R. J., & Amorim, R. D. M. (2023). O Inventário CBCL/6-18 no Brasil: Revisão de Evidências de Validade e Precisão. *Avaliação Psicológica*, 22(1), 33–41. <https://doi.org/10.15689/ap.2023.2201.21503.04>
- Rocha M. M., Rescorla, L. A., Emerich, D. R. Silveiras, E. F., Borsa, J. C., Araújo, L. G., Bertolla, M. H., Oliveira, M. S., Perez, N. C., Freitas, P. M., & Assis, S. G. (2013). Behavioural/emotional problems in Brazilian children: Findings from parents' reports on the Child Behavior Checklist. *Epidemiology and Psychiatric Sciences*, 22(4), 329–338. <https://doi.org/10.1017/S2045796012000637>
- Rodrigues, E. C., & Alchieri, J. C. (2009) Avaliação das características de afetividade em crianças e jovens com síndrome de Down. *Psico-USF*, 14(1), 107–116. https://pepsic.bvsalud.org/scielo.php?script=sci_abstract&pid=S1413-82712009000100011
- Rodríguez, E. R. (2004). Programa de educación emocional para niños y jóvenes com síndrome de Down. *Revista Síndrome de Down*, 21, 84-93. http://revistadown.downcantabria.com/wp-content/uploads/2004/09/revista82_84-93.pdf
- Sato, M. (1987). *Métodos de análise fatorial em psicologia*. EdUSP.
- Silveiras, E. F. M., Meyer, S. B., Santos, E. O., & Gerencer, T. T. (2006). Um estudo em cinco clínicas-escola brasileiras com a lista de verificação comportamental para criança (CBCL). In E. F. M., Silveiras (Org.), *Atendimento Psicológico em Clínicas-Escola* (pp. 59–72). Alínea.
- Taber, K. S. (2018). *The use of cronbach's alpha when developing and reporting research instruments in science education*. *Research in Science Education*, 48(6), 1273–1296. <https://link.springer.com/article/10.1007/s11165-016-9602-2>
- Urbina, S. (2014). *Essentials of psychological testing*. John Wiley & Sons.

Recebido em: 12/09/2024

Última revisão: 07/02/2025

Aceite final: 13/02/2025

Sobre os autores:

Rauni Jandé Roama-Alves: [Autor para contato]. Estágio pós-doutoral em Fonoaudiologia pela Universidade de São Paulo (USP). Doutor em Psicologia pela Pontifícia Universidade Católica de Campinas (PUC-Campinas). Docente do Programa de Pós-Graduação em Psicologia da Universidade Federal de Mato Grosso (UFMT). **E-mail:** rauniroama@gmail.com, **Orcid:** <https://orcid.org/0000-0002-1982-1488>

Rosimeire de Moraes Amorim: Mestre em Psicologia pela Universidade Federal de Mato Grosso. Especialização em Terapia Cognitiva Comportamental (Faculdade Poliensino). Docente do curso de Psicologia do Centro Universitário de Várzea Grande (Univag). **E-mail:** rosiamorim@gmail.com, **Orcid:** <https://orcid.org/0000-0003-2459-4053>

Tatiane Lebre Dias: Doutora em Psicologia pela Universidade Federal do Espírito Santo. Docente do Departamento de Psicologia e do Programa de Pós-Graduação em Psicologia da Universidade Federal de Mato Grosso. Pesquisadora do Laboratório de Estudos e Pesquisas em Avaliação Psicológica (UFMT). Membro do GT de Psicologia da Saúde da Criança e do Adolescente (Anpepp). **E-mail:** tatiane.dias@ufmt.br, **Orcid:** <https://orcid.org/0000-0002-9515-1578>

Rosângela Kátia Sanches Mazzorana Ribeiro: Estágio pós-doutoral em Psicologia Clínica na Universidade de São Paulo (USP). Doutora em Psiquiatria e Psicologia Médica pela Universidade Federal de São Paulo (Unifesp). Professora associada do curso de Psicologia e do Programa de Pós-Graduação em Psicologia da Universidade Federal de Mato Grosso (UFMT). **E-mail:** rosangelaksm@uol.com.br, **Orcid:** <https://orcid.org/0000-0003-4072-1091>

Raiane Mariah Roama Alves: Graduanda em Medicina pela Faculdade de Pinhais (Fapi). **E-mail:** raiane.mariah.3@gmail.com, **Orcid:** <https://orcid.org/0009-0000-4281-0077>

Carolina Rosa Campos: Doutora em Psicologia, mestre em Psicologia e psicóloga pela Pontifícia Universidade Católica de Campinas. Professora do Programa de Pós-Graduação em Psicologia da Universidade Federal do Triângulo Mineiro. **E-mail:** carolinarosacampos@gmail.com, **Orcid:** <https://orcid.org/0000-0002-1713-3307>