

Propriedades Psicométricas da Versão Brasileira do Battelle Developmental Inventory Screening: Crianças até Dois Anos

Psychometric Properties of the Brazilian Version of the Battelle Developmental Inventory Screening: Children up to Two Years Old

Propiedades Psicométricas de la Versión Brasileña del Battelle Developmental Inventory Screening: Niños de hasta Dos Años de Edad

Karolina Alves de Albuquerque

Universidade Federal do Espírito Santo (UFES)

Ana Cristina Barros da Cunha

Márcia Leonardi Baldisserotto

Pedro Paulo Pires

Universidade Federal do Rio de Janeiro (UFRJ)

Resumo

Introdução: O *Battelle Developmental Inventory Screening* foi adaptado para o Brasil (BDIS-BR), mas suas propriedades psicométricas ainda não foram testadas. O objetivo deste estudo foi investigar evidências psicométricas do BDIS-BR em 282 crianças menores de dois anos. **Métodos:** A validade baseada na estrutura interna foi verificada adotando-se Modelos de Mokken e validade convergente e concorrente foram verificadas por meio de tendências de idade e comparação com o Teste Denver II. **Resultados:** O BDIS-BR se mostrou como uma medida unidimensional e o ordenamento empírico dos itens tal qual o original. Quanto à validade convergente e concorrente, os resultados no BDIS-BR apresentam forte associação positiva com a idade das crianças ($p = 0,92$; $p = 0,001$) e altos índices de concordância com o Teste Denver II (86% a 96%). **Conclusão:** O BDIS-BR é um instrumento de triagem do desenvolvimento infantil com boas propriedades psicométricas e pode contribuir para a avaliação de crianças brasileiras.

Palavras-chave: estudos de avaliação, estudos de validação, desenvolvimento infantil

Abstract

Introduction: The *Battelle Developmental Inventory Screening* was adapted for Brazil (BDIS-BR), but its psychometric properties have not yet been tested. The aim of this study was to investigate the psychometric evidence of the BDIS-BR in 282 children under two years of age. **Methods:** The validity based on the internal structure was verified using Mokken Models, and convergent and concurrent validity were verified through age trends and comparison with the Denver II Test. **Results:** The BDIS-BR was shown to be a unidimensional measure and the empirical ordering of the items was similar to the original. Regarding convergent and concurrent validity, the results in the BDIS-BR showed a strong positive association with the age of the children ($p = 0.92$; $p = 0.001$), and high levels of agreement with the Denver II Test (86% to 96%). **Conclusion:** The BDIS-BR is a child development screening instrument with good psychometric properties and can contribute to the assessment of Brazilian children.

Keywords: evaluation studies, validation studies, child development

Resumen

Introducción: El *Battelle Developmental Inventory Screening* fue adaptado para Brasil (BDIS-BR), pero sus propiedades psicométricas aún no han sido probadas. El objetivo de este estudio fue investigar la evidencia psicométrica del BDIS-BR en 282 niños menores de dos años. **Métodos:** La validez basada en la estructura interna se verificó mediante Modelos de Mokken y la validez convergente y concurrente se verificó mediante tendencias de edad y comparación con el Test de Denver II. **Resultados:** El BDIS-BR resultó ser una medida unidimensional y el ordenamiento empírico de los ítems fue igual al original. En cuanto a la validez convergente y concurrente, los resultados del BDIS-BR muestran una fuerte asociación positiva con la edad de los niños ($p = 0,92$; $p = 0,001$) y altos niveles de concordancia con el Test Denver II (86% a 96%). **Conclusión:** El BDIS-BR es un instrumento de evaluación del desarrollo infantil con buenas propiedades psicométricas y puede contribuir a la evaluación de los niños brasileños.

Palabras clave: estudios de evaluación, estudios de validación, desarrollo infantil

Introdução

As experiências nos dois primeiros anos de vida são primordiais para o desenvolvimento motor, cognitivo, emocional, comportamental e social (Darling et al., 2020). A vigilância do desenvolvimento, com testes de triagem padronizados para identificação de atrasos, tem função preventiva importante e deve fazer parte das políticas públicas e dos programas de atenção ao desenvolvimento da primeira infância (Morris et al., 2020; Munoz-Chereau et al., 2021). Testes de triagem são recursos para uma avaliação breve que identificam crianças em risco para atrasos no desenvolvimento, global ou em áreas específicas, as quais devem ser encaminhadas para avaliação diagnóstica mais ampla (Richardson, 2010; Sabanathan et al., 2015). O uso de ferramentas sensíveis e com boas propriedades psicométricas para avaliação da criança é um grande desafio para profissionais e pesquisadores brasileiros (Albuquerque & Cunha, 2020).

O *Battelle Developmental Inventory, 2nd Edition* (BDI-2), é um instrumento padronizado de avaliação do desenvolvimento de crianças entre zero e sete anos e 11 meses (Newborg, 2005). O BDI-2 contém 450 itens e é organizado em cinco escalas chamadas “domínios”: Adaptativo, Pessoal-Social, Comunicativo, Motor e Cognitivo. Ele tem uma versão curta para triagem, o *Battelle Developmental Inventory Screening* (BDIS), com 100 itens, sendo dois itens por faixa etária para cada domínio. O BDIS foi considerado o instrumento de triagem com melhor sensibilidade e validade de critério para crianças menores de cinco anos de idade, nos EUA e na América Latina (Romo-Pardo et al., 2012), e tem sido usado em pesquisas em todo o mundo (Febrina et al., 2023; Peters & Matson, 2020; Prieto et al., 2022).

Uma versão traduzida e adaptada do BDI-2, e consequentemente do BDIS, para o Português-Brasil, foi obtida em um estudo de adaptação transcultural que envolveu a avaliação de uma versão brasileira do instrumento conduzida em rigoroso processo metodológico, em que foram avaliadas as equivalências semântica, idiomática, conceitual, experiencial e operacional, além da validação de conteúdo dos itens do instrumento para uso com a população infantil brasileira (Albuquerque et al., 2022). A versão adaptada para o português-Brasil do BDI-2, assim como sua versão de triagem denominada aqui como BDIS-BR, pode ser considerada uma nova ferramenta para profissionais das áreas de saúde e da educação do país. Todavia, são necessários estudos para a investigação das evidências de validade e de confiabilidade destas versões. Baseado na necessidade de se propor instrumentos válidos para programas de saúde e educação infantil e para a identificação precoce de crianças brasileiras com possíveis atrasos no desenvolvimento, o objetivo principal deste estudo foi investigar propriedades psicométricas do BDIS-BR para confirmar sua validade para uso com a população infantil brasileira até dois anos.

Método

Participantes

Uma amostra de 282 crianças com idade entre zero e 24 meses participou nesse estudo de investigação das propriedades psicométricas da versão do *Battelle Developmental Inventory Screening* para o Brasil: o BDIS-BR. As crianças foram recrutadas para o estudo de acordo com a seguinte classificação por cotas: sexo da criança, classe socioeconômica (crianças atendidas pelo SUS ou por planos privados) e idade em meses (0 a 5 meses, 6 a 11

meses, 12 a 17 meses, 18 a 24 meses). Foram incluídas crianças típicas e atípicas, sendo 240 crianças de uma amostra não clínica, 30 crianças com deficiências múltiplas e 12 crianças pré-termo. O projeto foi aprovado pelo Comitê de Ética da Universidade Federal do Espírito Santo (CAAE: 62650416.4.0000.5060).

Instrumentos

Protocolo de Dados Gerais e Critério de Classificação Econômica Brasil (CCEB)

Um protocolo de dados gerais foi elaborado para coletar dados clínicos das crianças e dados sociodemográficos das famílias. Os pais também responderam ao CCEB, um instrumento que identifica o potencial de consumo de famílias brasileiras e classifica o nível socioeconômico (NSE) baseado em 35 indicadores de renda (Associação Brasileira de Empresas de Pesquisa [ABEP], 2014).

Versão brasileira do Battelle Developmental Inventory Screening (BDIS-BR)

O BDIS-BR é uma medida de avaliação do desenvolvimento infantil adaptada para o Português-Brasil da versão *Screening do Battelle Developmental Inventory, 2nd Edition*—BDIS (Newborg, 2005). O BDIS é uma ferramenta de triagem do desenvolvimento, referenciada em normas e usada para identificar riscos para atraso no desenvolvimento em crianças de zero a sete anos e 11 meses de idade (Newborg, 2005). Ele avalia cinco domínios do desenvolvimento (Adaptativo, Pessoal-Social, Comunicação, Motor e Cognitivo) e provê um escore total e por domínios que são avaliados por 100 itens. Os itens são alocados em ordem de complexidade em cada um dos domínios. A administração do BDIS se inicia com a avaliação da criança pelos itens correspondentes à sua idade e continua até ela atingir seu nível “teto”, ou seja, uma pontuação zero em três itens consecutivos ou no item mais alto do domínio. Essa administração pode ser feita por observações da criança, entrevistas aos cuidadores e situação de interação estruturada. O desempenho da criança é classificado com base em um sistema simples de escores de três pontos, sendo 0 = incapaz, 1 = habilidade emergente e 2 = habilidade consolidada. Por fim, a pontuação obtida pela somatória dos itens até o nível “teto” gera um escore bruto total e por domínio que, comparados a uma tabela de pontos de corte (-1,0, -1,5 ou -2,0 desvio-padrão), determinam risco para atraso de desenvolvimento, podendo ser classificados como “normal” ou “suspeito” (Newborg, 2005).

Teste de Triagem do Desenvolvimento Denver II

O Teste Denver II é a ferramenta para rastreamento de riscos para atrasos do desenvolvimento infantil mais utilizado no Brasil (Lopez Boo et al., 2020). Ele pode ser usado em crianças de zero a seis anos de idade e apresenta 125 itens que avaliam o desenvolvimento infantil em quatro áreas: “Motor Grosso”, “Motor Fino-Adaptativo”, “Pessoal Social” e “Linguagem”. Propriedades psicométricas de uma versão traduzida para o Português foram recentemente testadas e apresentaram índices satisfatórios de sensibilidade e especificidade, variando de 73%-99% e 58%-92%, respectivamente (Santos et al., 2022). O desempenho da criança é identificado como “normal”; “suspeito” (um atraso); e “anormal” (dois ou mais atrasos). Para este estudo, “suspeito” e “anormal” foram considerados como uma única classificação.

Procedimentos

A coleta de dados foi feita por três examinadores, profissionais graduados em Psicologia ou Terapia Ocupacional, após rigoroso treinamento com base no estudo do manual, uso de vídeos e aulas teóricas e práticas sobre o BDI-2. A coleta de dados com as famílias durou cerca de dez minutos, enquanto as avaliações das crianças com o BDIS-BR tiveram duração média de 20 minutos. Para investigar a validade concorrente, 50 crianças, selecionadas de forma aleatória da amostra total, foram avaliadas pelo Teste Denver II. Ambos os testes, BDIS-BR e Denver II, foram administrados por diferentes examinadores, treinados e cegos ao resultado do outro instrumento.

Análise estatística

Estatísticas descritivas foram utilizadas para caracterizar o perfil clínico e sociodemográfico da população estudada e descrever o desempenho das crianças em termos de escore total e por domínios do BDIS-BR, por faixa etária de idade. As diferenças de sexo nos escores médios do BDIS-BR foram comparadas usando o teste *t de Student* com *Bootstrap*.

Considerando que o BDIS-BR é um instrumento de desempenho com itens escalonados de forma categórica ordinal, para averiguação de suas propriedades psicométricas, utilizaram-se os modelos de Mokken, pertencentes à Teoria de Resposta ao Item (TRI). Para essas análises, os seis primeiros itens de cada escala correspondentes à faixa etária estudada foram utilizados, totalizando 30 itens.

Através dos Modelos Mokken, obteve-se o coeficiente H de Loevinger para pares de itens (H_{ij}), para cada item individual (H_j) e para toda a escala (H). No caso de adequação e validade desses pressupostos, a Escala de Mokken foi classificada como fraca ($0,30 \leq H < 0,40$), razoável ($0,40 \leq H < 0,50$) ou forte ($H \geq 0,50$) (Van der Ark, 2012). Também foi empregado o algoritmo *Automated Item Selection Procedure* (AISP), que considera dois parâmetros para seleção de itens a serem mantidos ou não no modelo testado: (1) o limite inferior da covariância inter-item precisa ser superior a zero e positiva e (2) significativa. O algoritmo foi empregado para fins de confirmação das interpretações realizadas.

Para averiguar evidências de validade convergente, adotou-se o teste de correlação de *Spearman* entre os escores brutos do BDIS-BR e a idade em meses das 282 crianças. Correlações $\pm 0,30$ são consideradas fracas, $\pm 0,50$ moderadas e $\pm 0,70$ são consideradas fortes (Rumsey, 2023). Além disso, foram plotadas curvas de tendência de idade em meses e escores em cada domínio do BDIS-BR. Porcentagens foram calculadas para analisar a concordância entre esses resultados assumidos como variáveis categóricas: com e sem risco de atraso de desenvolvimento para os resultados do BDIS-BR e do Teste de Denver II.

As análises descritivas e de validade convergente foram feitas no programa *Statistical Package for Social Science* (SPSS), versão 23.0, adotando-se um nível de significância $\alpha = 0,05$. O gráfico de curvas de tendência polinomial foi plotado no programa Excel for Windows. A análise das Escalas de Mokken foi realizada pelo Programa R., interface t R Studio, versão 3.5.1. (pacote *Mokken*).

Resultados

A amostra foi composta por 282 crianças, 143 meninos e 139 meninas. Não houve diferença entre meninos e meninas em relação ao desempenho no BDIS-BR ($t=-0,06$; $g/= 280$; $p=0,93$). A idade média foi de 13,49 meses ($DP=6,27$) (Tabela 1).

Tabela 1

Dados Clínicos e Sociodemográficos das Crianças e suas Famílias e o Desempenho das Crianças no BDIS-BR

Variáveis	0 – 5 meses (n= 47)	6 – 11 meses (n=54)	12 – 17 meses (n=89)	18 – 23 meses (n=92)	Total (N=282)
Idade ^a	3,68(1,64)	8,52 (1,8)	14,7 (1,84)	20,25 (1,73)	13,49 (6,27)
Sexo ^b					
Masculino	25 (53,2%)	24 (44,4%)	46 (51,7%)	48 (52,2%)	143 (50,7%)
Feminino	22 (46,8%)	30 (55,6%)	43 (48,3%)	44 (47,8%)	139 (49,3%)
Idade gestacional (semanas) ^a	38,6 (2,32)	37,59 (3,52)	38,44 (2,63)	38,41 (3,07)	38,29 (2,92)
Peso ao nascer (gramas) ^a	3161,51 (705,75)	2961,04 (781,14)	3128,09 (671,44)	3183,23 (717,29)	3119,43 (714,75)
Creche ^b					
Sim	1 (2,3%)	10 (20,8%)	34 (39,5%)	58 (63,7%)	103 (38,3%)
Não	43 (97,7%)	38 (79,2%)	52 (60,5%)	33 (36,3%)	166 (61,7%)
NSE ^b					
A	7 (14,9%)	5 (9,3%)	19 (21,3%)	16 (17,4%)	47 (16,7%)
B	16 (34,1%)	25 (46,3%)	39 (43,8%)	37 (40,3%)	117 (41,5%)
C	14 (29,8%)	13 (24,1%)	18 (20,2%)	25 (27,2%)	70 (24,8%)
D / E	6 (12,8)	2 (3,7%)	9 (10,1%)	13 (14,1%)	30 (10,6%)
Não informado	4 (8,5%)	9 (16,7%)	4 (4,5%)	1 (1,1%)	18 (6,4%)
Idade materna (anos) ^a	31,09 (6,52)	30,45 (6,48)	32,16 (5,98)	30,86 (5,64)	31,23 (6,07)
Escolaridade materna (anos) ^a	13,19 (3,77)	13,30 (3,76)	13,67 (3,60)	13,25 (3,96)	13,38 (3,77)
Doença psiquiátrica materna ^b					
Sim	8 (18,2%)	10 (20,8%)	19 (22,1%)	12 (13%)	49 (18,1%)
Não	36 (81,8%)	38 (79,2%)	67 (77,9%)	80 (87%)	221 (81,9%)
BDIS-BR – ADP ^a	3,77 (1,46)	7,61 (2,10)	12,70 (3,83)	15,88 (4,74)	11,27 (5,71)
BDIS-BR – P-S ^a	6,42 (2,26)	10,35 (2,53)	16,30 (3,92)	19,03 (4,90)	14,41 (6,05)
BDIS-BR – COM ^a	4,68 (1,59)	8,68 (2,83)	13,12 (3,43)	15,94 (4,56)	11,78 (5,37)
BDIS-BR – MOT ^a	4,11 (1,98)	8,57 (2,72)	13,91 (3,93)	16,27 (5,45)	12,02 (6,02)
BDIS-BR – COG ^a	6,28 (1,98)	10,80 (3,27)	14,48 (3,24)	15,91 (4,32)	12,87 (4,89)
BDIS-BR – Total ^a	25,25 (7,94)	46,02 (11,66)	70,53 (15,71)	83,04 (22,12)	62,37 (26,71)

Nota. ^aMédia (desvio-padrão); ^bFrequência (porcentagem); BDIS-BR: *Battelle Developmental Inventory Screening*, versão brasileira; ADP: domínio adaptativo; P-S: domínio pessoal-social; COM: domínio comunicativo; MOT: domínio motor; COG: domínio cognitivo.

Validade Baseada na Estrutura Interna – Escala de Mokken

A investigação do pressuposto de monotonicidade dos itens em relação à escala indicou que não houve violações em nenhum dos itens pertencentes às escalas (domínios) do BDIS-BR. O mesmo foi verificado para os pressupostos de item-total, cuja matriz p indica não haver problemas de interseção entre as ISRFs (*Item Step Response Function*) e nem para a IIO (*Invariant Item Ordering*). Tais dados atestam que cada item funciona de forma ascendente para o escore, ou seja, que eles contribuem linearmente para a composição do escore da escala (domínio) do instrumento.

Os coeficientes H de Loevinger para os seis itens de cada domínio, do BDIS-BR correspondentes à faixa etária, estão apresentados na Tabela 2. O item 1 do domínio Pessoal-Social apresentou o menor valor para o coeficiente H individual dos itens: $H_j=0,61$ ($EP=0,13$), ainda consideravelmente superior ao índice necessário (0,30) (Van der Ark, 2012). Em relação à análise para os pares de itens, a matriz do coeficiente H não apresentou problemas de ajuste local, com valores de H_{ij} superiores a 0,30. O coeficiente H de Loevinger da escala geral foi 0.82 (0.02), indicando qualidade psicométrica do Inventário.

Tabela 2

Coeficientes H de Loevinger para Cada Item das Escalas do BDIS-BR na Faixa Etária Estudada

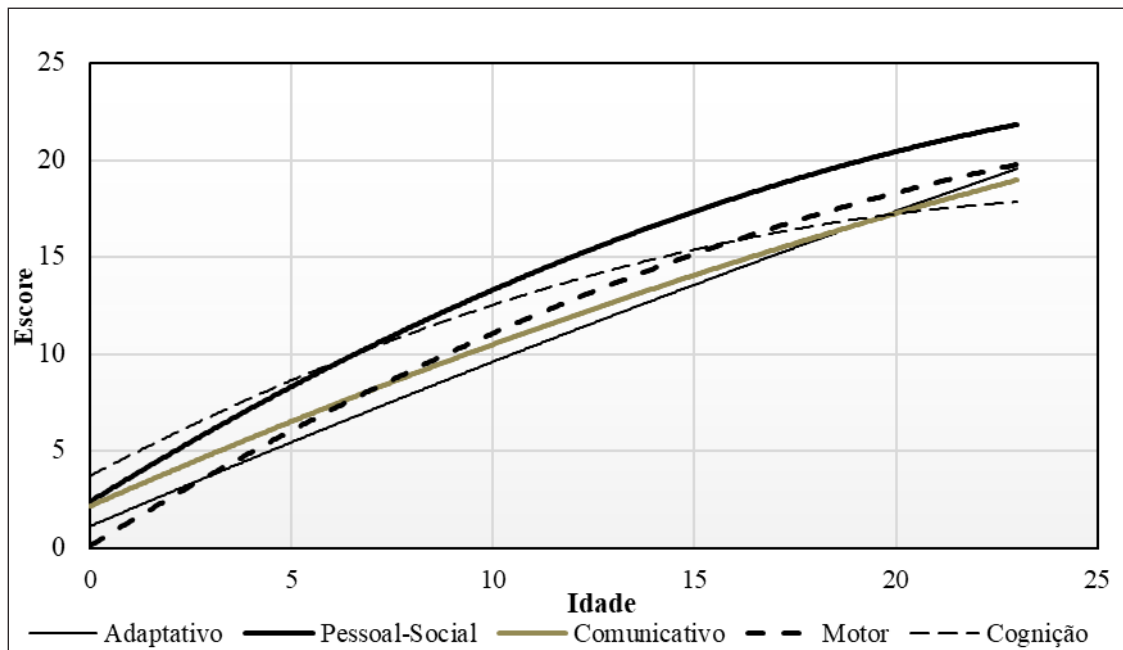
	Adaptativo		Pessoal-Social		Comunicativo		Motor		Cognitivo	
	Hi	Se	Hi	Se	Hi	Se	Hi	Se	Hi	Se
Item 1	0.62	0.06	0.61	0.13	0.63	0.07	0.78	0.04	0.65	0.10
Item 2	0.75	0.04	0.72	0.07	0.69	0.05	0.71	0.06	0.74	0.04
Item 3	0.83	0.03	0.66	0.06	0.72	0.04	0.87	0.02	0.77	0.04
Item 4	0.81	0.03	0.84	0.02	0.85	0.02	0.87	0.02	0.86	0.02
Item 5	0.87	0.02	0.78	0.03	0.83	0.03	0.86	0.02	0.74	0.03
Item 6	0.90	0.01	0.84	0.02	0.87	0.02	0.87	0.02	0.87	0.02

Validade Convergente – Tendências de Idade

A validade convergente foi investigada testando-se correlações bivariadas entre a idade da criança e seus resultados no BDIS-BR em termos de escore bruto. Correlações positivas fortes foram encontradas entre a idade da criança e o escore total do BDIS-BR ($p=0,92$; $p=0,001$), bem como todos os escores por domínios, sendo Adaptativo ($p=0,90$; $p=0,001$), Pessoal-Social ($p=0,85$; $p=0,001$), Comunicativo ($p=0,87$; $p=0,001$), Motor ($p=0,91$; $p=0,001$) e Cognitivo ($p=0,81$; $p=0,001$). O Gráfico 1 mostra uma série de cinco curvas de tendência dos cinco domínios do BDIS-BR de acordo com a idade da criança, que mostram que o desempenho da criança melhora à medida que sua idade aumenta.

Gráfico 1

Curvas de Tendência de Idade para os Domínios do BDIS-BR

**Validade Concorrente – Concordância com o Teste Denver II**

O BDIS-BR e o Teste Denver II foram administrados em uma amostra aleatória de 50 crianças do total da amostra, sendo 60% meninos. As porcentagens de concordância na classificação do desempenho da criança (normal e suspeito) no BDIS-BR e no Teste Denver II estão na Tabela 3. Essa porcentagem variou entre 86% e 96%, o que indica alto nível de concordância entre a classificação do BDIS-BR e do Denver II.

Tabela 3

Porcentagem de Concordância entre os Resultados das Crianças no BDIS-BR e no Teste de Denver II

BDIS-BR	Denver II			
	Pessoal-Social	Motor fino	Linguagem	Motor Grosso
Adaptativo	90%	86%		
Pessoal-Social	92%			
Comunicativo			96%	
Motor		90%		88%
Cognitivo			90%	

Nota. n=50; BDIS-BR: *Battelle Developmental Inventory Screening*, versão brasileira.

Discussão

Considerando que ferramentas de triagem do desenvolvimento infantil precisam apresentar integridade psicométrica para a identificação precoce de atrasos em crianças, o propósito deste estudo foi conduzir uma primeira avaliação das propriedades psicométricas da

versão brasileira do *Battelle Developmental Inventory Screening* (BDIS-BR). A condição *sine qua non* de qualquer ferramenta de avaliação para fins clínicos ou de pesquisa é apresentar boas propriedades psicométricas, o que foi observado nas análises de validade do BDIS-BR. Fortes evidências de integridade psicométrica do BDIS-BR foram encontradas, incluindo bons índices de validade da estrutura interna, convergente e concorrente. Isso sugere que o BDIS-BR pode ser considerado uma medida válida para uso com crianças brasileiras menores de dois anos.

O desempenho das crianças na avaliação pelo BDIS-BR mostrou que a média dos escores de todos os domínios está consideravelmente acima dos pontos de corte estabelecidos pelo instrumento original. Contanto, considerando o ponto de corte do BDIS original abaixo de 1,5 desvio-padrão, 9,57% da amostra apresentou suspeita de atraso do desenvolvimento global, porcentagem bastante próxima ao número de pessoas com deficiência no Brasil, de acordo com a Pesquisa Nacional por Amostra de Domicílio (PNAD) Contínua de 2022, cerca de 18,6 milhões de pessoas com deficiência, o que corresponde a 8,9% da população (IBGE, 2023). De modo geral, isso indica que foi possível identificar os riscos de atraso pela avaliação do escore geral do BDIS-BR, como proposto pelo instrumento. Além disso, não houve diferença de desempenho entre os sexos nos domínios do BDIS-BR, o que mostra que os itens da versão brasileira não apresentam viés de análise por sexo.

Ressalta-se que os dados da amostra deste estudo apresentaram baixa variabilidade de respostas para os itens iniciais instrumento, o que resultou em uma distribuição de dados não normal e no uso de um modelo não paramétrico da Teoria de Resposta ao Item (TRI), no caso as Escalas de Mokken. Essa baixa variabilidade ocorreu devido à grande maioria das crianças terem pontuado dois (2 = habilidade consolidada), enquanto poucas crianças pontuaram zero (0 = incapaz) ou um (1 = habilidade emergente) nos itens iniciais, que são os mais fáceis. Como o BDIS-BR é um teste de triagem, esse dado já era esperado, pois se trata de um tipo de teste que serve para o rastreamento do desenvolvimento infantil avaliado em poucos itens, porém suficientes para identificar parcialmente o traço latente correspondente ao constructo avaliado (Albuquerque & Cunha, 2020; Richardson, 2010; Sabanathan et al., 2015). Adicionalmente, o BDIS é um instrumento amigável e com flexibilidade processual, o que facilita à criança estar confortável na situação de testagem e ao examinador obter informações de diferentes fontes, o que pode resultar em um desempenho mais competente da criança, tal como afirmam Macy e colaboradores (Macy et al., 2015). Atributos como esse, junto da densidade dos itens, sua pontuação graduada e seu conteúdo funcional, conferem ao BDIS-BR o caráter de uma medida de avaliação apropriada, mesmo para crianças muito jovens. Estudos mostram que essas características aumentam a validade social do BDI-2 (Lee et al., 2016; Macy et al., 2015) e, consequentemente, pode sugerir que as versões adaptadas com boas propriedades psicométricas, como é o caso do BDIS-BR, também apresentam alta aceitabilidade e utilidade.

Para confirmar a estrutura do BDIS-BR, foi adotado um modelo de Teoria de Resposta ao Item (TRI) não paramétrico: a Escala de Mokken. O conjunto de itens dos domínios do BDIS-BR mostrou-se como uma medida unidimensional em uma escala ascendente (Baghaei, 2021), tal como no instrumento original. Adicionalmente, o ordenamento empírico dos itens correspondeu à ordem esperada, sustentando a validade das sequências das habilidades de desenvolvimento avaliadas pela escala de cada domínio do BDIS-BR. Ressalta-se a relevância

desse dado, já que o instrumento original se sustenta nos pressupostos dos marcos de desenvolvimento (Newborg, 2005), para o qual o desenvolvimento infantil resulta da aquisição de habilidades e capacidades de desenvolvimento que aumentam desde o nascimento (Papalia & Martorell, 2022).

A ordem dos itens para todos os domínios do BDIS-BR ficou idêntica à ordem prescrita de apresentação dos itens no instrumento original e em acordo com as análises originais conduzidas no estudo de Newborg usando TRI (Newborg, 2005). No estudo de Elbaum e colaboradores (Elbaum et al., 2010), os itens do teste de triagem do BDI-2 (BDIS) também foram submetidos a TRI usando o modelo de *Rasch*, quando ótimos ajustes dos itens ao modelo foram encontrados, de forma semelhante ao encontrado nas evidências de validade do BDIS-BR. Além disso, para todos os pares de itens, os coeficientes H de Loevinger foram satisfatórios (83% acima de 0,80), e isso reafirma a validade da estrutura interna e da equivalência da versão brasileira do BDIS com o instrumento original. Coeficientes dessa magnitude suportam a conclusão de que os itens estão relacionados ao construto medido pelas escalas (domínios) do BDIS-BR, que, portanto, representa uma medida apropriada para avaliar o desenvolvimento infantil também de crianças brasileiras.

Considerando que o BDIS-BR adota o pressuposto dos marcos do desenvolvimento, espera-se que as pontuações das escalas aumentem de acordo com a idade da criança (Papalia & Martorell, 2022). Assim, as fortes correlações positivas (entre 0,81 e 0,92) entre os escores, totais e por domínios, do BDIS-BR e a idade da criança representam evidências de validade desse instrumento para medir o construto, tal como pressuposto teoricamente no instrumento original. Similar ao estudo de padronização do BDI-2 (Newborg, 2005), as curvas de tendência da idade, observadas no Gráfico 1, reafirmam o aumento dos escores de desempenho das crianças ao longo do seu desenvolvimento até 24 meses de idade. De fato, nesse período da infância, espera-se um desenvolvimento exponencial e não linear em todos os domínios de desenvolvimento, o que é notável nas pontuações do BDIS-BR.

Para análise da validade concorrente do BDIS-BR, foi adotado o Teste Denver II, instrumento de triagem mais utilizado no Brasil e recomendado pela Sociedade Brasileira de Pediatria (Albuquerque & Cunha, 2020; Lopez Boo et al., 2020). Os índices de concordância da análise dos resultados do desempenho no BDIS-BR e no Teste Denver II para a população estudada foram altos: 86% a 96%. No estudo original de padronização do BDI-2 (Newborg, 2005), a validade concorrente também foi verificada pela concordância entre os resultados do BDI-2 e do Teste Denver II, em uma amostra de 72 crianças, e variou entre 83% e 90%, o que corrobora os resultados deste trabalho.

Comparados a outros estudos de validação psicométrica de instrumentos de triagem para o desenvolvimento infantil de crianças brasileiras, as evidências de validade encontradas para o BDIS-BR neste estudo são excelentes. Estudo de validação para o Brasil de outro instrumento de triagem do desenvolvimento infantil, o *Age and Stages Questionnaire, 3rd edition* (ASQ-3), também encontrou bons índices de validade baseado na estrutura interna, utilizando a Teoria Clássica de Testes (TCT) e a TRI com base nos Modelos *Rasch*. No entanto, pela análise baseada na TRI, 44 dos 540 itens do ASQ-3 apresentaram problemas de ajuste. Diferentemente do nosso estudo, o estudo de validação do ASQ-3 não investigou validades concorrentes ou convergentes (Filgueiras et al., 2013). Madaschi et al. (2016) também avaliaram propriedades psicométricas de outra ferramenta de avaliação infantil

muito usada, as escalas Bayley III. O estudo apresentou bons resultados nas análises fatoriais de consistência interna e validade concorrente, mas não foram feitas análises mais robustas, tais como TRI.

No geral, esta investigação mostra evidências de que as escalas (domínios) do BDIS-BR têm boas propriedades psicométricas e fornecem suporte para a aplicabilidade desse instrumento na população infantil brasileira. De todo modo, algumas limitações do estudo devem ser consideradas. Embora a amostra tenha sido composta com um número adequado de sujeitos, com distribuição equivalente entre sexo e idades, o critério “50% das crianças atendidas pelo somente pelo SUS” não foi suficiente para garantir a heterogeneidade de NSE (65% da amostra pertencia aos níveis A e B) (ABEP, 2014). Esse dado, relacionado aos dados familiares como idade ($M=31,58$ anos) e nível de escolaridade materna (59,58% ≥ 12 anos), sugerem que o perfil das crianças participantes representava uma parcela específica da população, o que dificulta a generalização dos dados. Tampouco foi possível incluir crianças de diferentes regiões do país, devido à dificuldade de estabelecer parcerias para uma coleta multicêntrica, o que é uma questão importante para estudos futuros.

Reconhecendo essas limitações, podemos concluir que o BDIS-BR se mostrou um instrumento adequado para rastrear indicadores de risco ao desenvolvimento de crianças brasileiras. As evidências psicométricas encontradas indicam que ele é um instrumento válido e confiável para identificar precocemente riscos de atrasos no desenvolvimento da população infantil menor de dois anos de idade. Testes de triagem como o BDIS-BR são alternativas viáveis e confiáveis para auxiliar no desafio de mensurar resultados do desenvolvimento infantil para crianças muito pequenas. Os resultados de uma ferramenta como o BDIS-BR podem ajudar no processo de triagem infantil e no encaminhamento de crianças em risco para serviços de saúde em que existam profissionais treinados para uma avaliação diagnóstica adicional completa. Esse é o caminho para que a prevenção de problemas de desenvolvimento possa ocorrer adotando-se diferentes formas de intervenções planejadas com antecedência. Por fim, ferramentas de triagem como o BDIS-BR podem melhorar a qualidade dos programas de vigilância do desenvolvimento, porque favorecem a tomada de decisão para intervenção precoce infantil e a aplicação dos investimentos governamentais, políticos e econômicos, para a promoção da saúde e do desenvolvimento na primeira infância.

Referências

- Associação Brasileira de Empresas de Pesquisa. (2014). *Critério de Classificação Econômica Brasil 2015*(p.1–6).
- Albuquerque, K. A. de, & Cunha, A. C. B. (2020). New trends in instruments for child development screening in Brazil: A systematic review. *Journal of Human Growth and Development*, 30(2), 188–196. <https://doi.org/10.7322/jhgd.v30.10366>
- Albuquerque, K. A., Cunha, A. C. B., & Berkovits, M. D. (2022). Cross-cultural adaptation of the Battelle Development Inventory, for Brazil. *Estudos de Psicologia (Campinas)*, 39, e190170. <https://doi.org/10.1590/1982-0275202239e190170>
- Baghaei, P. (2021). *Mokken Scale analysis in language assessment*. Waxmann Verlag GmbH.
- Darling, J. C., Bamidis, P. D., Burberry, J., & Rudolf, M. C. (2020). The First Thousand Days: Early, integrated and evidence-based approaches to improving child health –Coming to a population near you? *Archives of disease in childhood*, 105(9), 837–841.

- Elbaum, B., Gattamorta, K. A., & Penfield, R. D. (2010). Evaluation of the Battelle Developmental Inventory, 2nd Edition, Screening Test for Use in States' Child Outcomes Measurement Systems Under the Individuals With Disabilities Education Act. *Journal of Early Intervention*, 32(4), 255–274.
- Febrina, F., Lawalata, V. C. P., & Ramli, Y. (2023). Factors associated with cognitive and communication delay in children aged 0 to 3 years using the Battelle-Developmental Inventory, 2nd edition. *Paediatric a Indonesiana*, 63(4), 282–289. <https://doi.org/10.14238/pi63.4.2023.282-9>
- Filgueiras, A., Pires, P., Maissonette, S., & Landeira-Fernandez, J. (2013). Psychometric properties of the Brazilian-adapted version of the Ages and Stages Questionnaire in public child daycare centers. *Early Human Development*, 89(8), 561–576. <https://doi.org/10.1016/j.earlhumdev.2013.02.005>
- Instituto Brasileiro de Geografia e Estatística. (2023). *Pessoas com deficiência 2022: PNAD contínua*. https://biblioteca.ibge.gov.br/visualizacao/livros/liv102013_informativo.pdf
- Lee, D., Bagnato, S., & Pretti-Frontczak, K. (2016). Utility and Validity of Authentic Assessments and Conventional Tests for International Early Childhood Intervention Purposes: Evidence from U.S. National Social Validity Research. *Journal of Intellectual Disability – Diagnosis and Treatment*, 3(4), 164–176. <https://doi.org/10.6000/2292-2598.2015.03.04.2>
- Lopez Boo, F., Cubides Mateus, M., & Llonch Sabatés, A. (2020). Initial psychometric properties of the Denver II in a sample from Northeast Brazil. *Infant Behavior and Development*, 58, 101391. <https://doi.org/10.1016/j.infbeh.2019.101391>
- Macy, M., Bagnato, S. J., Macy, R. S., & Salaway, J. (2015). Conventional Tests and Testing for Early Intervention Eligibility: Is There an Evidence Base? *Infants and Young Children*, 28(2), 182–204. <https://doi.org/10.1097/IYC.0000000000000032>
- Madaschi, V., Mecca, T. P., Macedo, E. C., & Paula, C. S. (2016). Bayley-III Scales of Infant and Toddler Development: Transcultural Adaptation and Psychometric Properties. *Paidéia (Ribeirão Preto)*, 26(64), 189–197. <https://doi.org/10.1590/1982-43272664201606>
- Morris, A. S., Wakschlag, L., Krogh-Jespersen, S., Fox, N., Planalp, B., Perlman, S. B., Shuffrey, L. C., Smith, B., Lorenzo, N. E., Amso, D., Coles, C. D., & Johnson, S. P. (2020). Principles for Guiding the Selection of Early Childhood Neurodevelopmental Risk and Resilience Measures: HEALTHy Brain and Child Development Study as an Exemplar. *Adversity and Resilience Science*, 1(4), 247–267. <https://doi.org/10.1007/s42844-020-00025-3>
- Munoz-Chereau, B., Ang, L., Dockrell, J., Outhwaite, L., & Heffernan, C. (2021). Measuring early child development across low and middle-income countries: A systematic review. *Journal of Early Childhood Research*, 19(4), 443–470. <https://doi.org/10.1177/1476718X211020031>
- Newborg, J. (2005). *Battelle Developmental Inventory, 2nd edition, Examiner's Manual* (2^o ed.). Riverside Publishing.
- Papalia, D. E; Martorell, G. (2022). *Desenvolvimento humano* (14^aed). AMGH.
- Peters, W. J., & Matson, J. L. (2020). The Relationship Between Developmental Functioning and Screening Outcome for Autism Spectrum Disorder. *Journal of Developmental and Physical Disabilities*, 32(2), 293–305. <https://doi.org/10.1007/s10882-019-09689-x>

- Prieto, J. A., Cueto, S., Carballo-Fazanes, A., & Abelairas, C. (2022). Psychomotor development disorders in apparently healthy children and considerations of family evaluation. *Journal of Human Sport and Exercise*. 17(3), 518–527. <https://doi.org/10.14198/jhse.2022.173.04>
- Richardson, K. P. (2010). Use of standardized tests in pediatric practice. In *Occupational Therapy for Children* (p. 216–243). Mosby.
- Romo-Pardo, B., Liendo-Vallejos, S., Vargas-López, G., Rizzoli-Córdoba, A., & Buenrostro-Márquez, G. (2012). Pruebas de tamizaje de neurodesarrollo global para niños menores de 5 años de edad validadas en Estados Unidos y Latinoamérica: Revisión sistemática y análisis comparativo. *Boletín Médico del Hospital Infantil de México*, 69(6), 450–462.
- Rumsey, D. J. (2023). What is r value correlation? *Dummies*. <https://www.dummies.com/article/academics-the-arts/math/statistics/how-to-interpret-a-correlation-coefficient-r-169792/>
- Sabanathan, S., Wills, B., & Gladstone, M. (2015). Child development assessment tools in low-income and middle-income countries: How can we use them more appropriately? *Archives of Disease in Childhood*, 100(5), 482–488. <https://doi.org/10.1136/archdischild-2014-308114>
- Santos, J. A. T., Ayupe, K. M. A., Lima, A. L. O., Albuquerque, K. A. de, Morgado, F. F. da R., & Gutierrez Filho, P. J. B. (2022). Propriedades psicométricas da versão brasileira do Denver II: Teste de triagem do desenvolvimento. *Ciência & Saúde Coletiva*, 27, 1097–1106. <https://doi.org/10.1590/1413-81232022273.40092020>
- Van Der Ark, L. A. (2012). New Developments in Mokken Scale Analysis in R. *Journal of Statistical Software*, 48(5), 1–27. <https://doi.org/10.18637/jss.v048.i05>

Recebido em: 13/09/2024

Última revisão: 12/01/2025

Aceite final: 12/01/2025

Sobre os autores:

Karolina Alves de Albuquerque: [Autora para contato]. Pós-doutorado em Psicologia pela Universidade Federal do Rio de Janeiro (UFRJ). Doutora em Psicologia pela Universidade Federal do Espírito Santo (UFES). Docente do Curso de Graduação em Terapia Ocupacional na UFES. **E-mail:** karolina.albuquerque@ufes.br, **Orcid:** <https://orcid.org/0000-0002-6044-3570>

Ana Cristina Barros da Cunha: Pós-doutorado pela University of Miami, EUA. Doutorado em Psicologia Social e do Desenvolvimento pela Universidade Federal do Espírito Santo (UFES). Professora titular na Universidade Federal do Rio de Janeiro (UFRJ). Docente e pesquisadora de Psicologia Perinatal, do Desenvolvimento e da Parentalidade, no Instituto de Psicologia/Maternidade Escola, UFRJ. **E-mail:** acbcunha@yahoo.com.br, **Orcid:** <https://orcid.org/0000-0003-0839-0130>

Márcia Leonardi Baldisserotto: Doutora em Epidemiologia em Saúde Pública pela Escola Nacional de Saúde Pública Sérgio Arouca (ENSP), vinculada à Fundação Oswaldo Cruz (FIOCRUZ). Professora adjunta no Instituto de Psicologia da Universidade Federal do Rio de Janeiro (UFRJ). **E-mail:** mlbaldisserotto@gmail.com, **Orcid:** <https://orcid.org/0000-0001-6907-2510>

Pedro Paulo Pires: Doutor em Psicologia pela Universidade Federal do Rio de Janeiro (UFRJ). Professor e supervisor clínico pelo Depto. de Psicometria (UFRJ). **E-mail:** ppires85@gmail.com, **Orcid:** <https://orcid.org/0000-0001-8831-9056>