

Evidências de validade baseada nas relações com medidas externas e confiabilidade do IFIS - BRA

<https://doi.org/10.11606/issn.1981-4690.2024e38204419>

Hildeberto Dutra Sobral*
Jânio Luiz Correia Júnior*
Débora de Almeida Pereira*
Marina Pereira Gonçalves**
Hamilton Felipe Andrade Santos***
Nelson Carvas Junior****
Ricardo de Freitas Dias*/****

*Universidade de Pernambuco, Programa de Pós-graduação em Hebiatria, Recife, PE, Brasil.

**Universidade Federal do Vale do São Francisco, Petrolina, PE, Brasil.

***Universidade de Pernambuco, Petrolina, PE, Brasil.

****Universidade Ibirapuera, Departamento de Fisioterapia, São Paulo, SP, Brasil.

Resumo

A avaliação do nível da aptidão física em adolescentes apresenta-se como preditor de doenças crônicas degenerativas da fase adulta, logo a sua mensuração por meio simples, barato e de fácil aplicabilidade faz com que o International Fitness Scale, seja de interesse da saúde pública mundial. Portanto, nosso objetivo foi avaliar as evidências baseadas nas relações com variáveis externas e confiabilidade do International Fitness Scale - versão brasileira (IFIS - BRA). Participaram do estudo adolescentes, de ambos os sexos. Foram aplicados o IFIS - BRA e testes físicos nos adolescentes. Tomados em conjunto os resultados, podemos afirmar que o IFIS - BRA, oferece indicadores psicométricos satisfatórios.

PALAVRAS-CHAVE: Aptidão física; Adolescente; Questionário; Estudo de validação.

Introdução

A aptidão física é a capacidade de realizar atividades físicas, de maneira satisfatória, possuindo como componentes: força muscular (FM), flexibilidade (Flex), capacidade cardiorrespiratória (CC), velocidade / agilidade (VA)¹. A avaliação do nível da aptidão física em adolescentes é de interesse da saúde pública mundial, por se apresentar como um preditor de doenças crônicas degenerativas da fase adulta. Estima-se que, a obesidade causou 24.678 óbitos de indivíduos, no Brasil, durante o período de 2010 a 2021². Com o crescente aumento de adolescentes com sobrepeso e obesidade

torna-se fundamental avaliar a aptidão física para prevenção de doenças crônicas³.

A utilização de testes físicos considerados “padrão-ouro”, para avaliar a aptidão física, geralmente, é de alto custo e necessita de especialistas, além de levar muito tempo para sua execução e possuir dificuldades na avaliação de grandes grupos de adolescentes ao mesmo tempo, tornando-se, às vezes, seu uso limitante em locais que não possuem essas propriedades⁴. O desenvolvimento de um instrumento autorrelatado preciso para avaliar a aptidão física, pode ser utilizado para identificar indivíduos com baixa

aptidão física e direcionar condutas, de forma simples e prática⁵.

Assim, surge a necessidade de buscar métodos alternativos, mais baratos e acessíveis, como os questionários de autorrelato⁵⁻⁹. Contudo, para realizar a mensuração dos resultados são necessários instrumentos que tenham alta acurácia ou consistência, eficácia em prever o desempenho específico de um sujeito (validade de critério) e capacidade de reproduzir os resultados de forma semelhante, no tempo e espaço (confiabilidade), sendo um dos principais requisitos a padronização das instruções dos questionários¹⁰.

Em virtude da necessidade de mensurar aptidão física por meios alternativos, surge o *International*

Fitness Scale (IFIS), questionário que avalia aptidão física geral (AFG) e seus domínios, CC, FM, VA e Flex⁵. O questionário foi validado e adaptado culturalmente para o idioma português do Brasil e apresentou boa validade baseada no conteúdo para adolescentes^{11,12}. No entanto, para indicar se o questionário IFIS, em sua versão português brasileiro (IFIS-BRA), apresenta capacidade de prever um comportamento futuro relacionado à aptidão física e seu erro de medição, este estudo teve como objetivo avaliar as evidências de validade baseada nas relações com medidas externas (objetivo 1) e confiabilidade (objetivo 2) do IFIS-BRA.

Método

Aspectos éticos

Este estudo atendeu aos critérios de ética em pesquisas com seres humanos, visando respeitar a Resolução n.º 466/12, do Conselho Nacional de Saúde. Foi aprovado pelo Comitê de Ética em Pesquisa (CEP), da Universidade de Pernambuco (UPE) sob o n.º do parecer 3.015.407 e CAAE 93550718.0.0000.5207. Antes do início da pesquisa, todos os participantes tiveram autorização dos pais ou responsáveis e entregaram o Termo de Consentimento Livre e Esclarecido (TCLE) e o Termo de Assentimento Livre e Esclarecido (TALE) assinados.

Amostra

Para a validade de critério do tipo concorrente (objetivo 1), foram selecionados, por conveniência, 164 adolescentes e para avaliar a confiabilidade (objetivo 2), uma nova amostra de 104 adolescentes, de ambos os sexos. Como critério de inclusão, os adolescentes deveriam ter faixa etária entre 12 e 17 anos. Como critérios de exclusão, foram considerados os participantes que não preencheram completamente o questionário IFIS-BRA, que apresentaram comprometimento das condições de saúde física ou intelectual, declarado previamente pelo professor, e gestantes.

Procedimentos

Inicialmente, o avaliador passou por um treinamento de calibração e padronização das instruções para aplicação e coleta dos dados. Em seguida foram realizadas as coletas dos dados dos objetivos 1 e 2.

Validade de critério (objetivo 1)

Para a validade de critério do tipo concorrente foi realizado a aplicação do questionário IFIS-BRA, seguido da realização das avaliações: antropométrica, CC, FM, VA e Flex, com intervalos de 24 a 48 horas entre os testes físicos.

Confiabilidade (objetivo 2)

Inicialmente, foi realizada a padronização das instruções para a aplicação do IFIS-BRA, e avaliada por dois avaliadores independentes. Ao final de um encontro, buscaram o consenso e a obtenção de uma única versão^{12,13}. Todas as discrepâncias que não puderam ser resolvidas por consenso, entre os dois avaliadores, foram julgadas por um terceiro avaliador. Em seguida, os adolescentes foram divididos, em dois grupos: grupo que respondeu o IFIS-BRA, sem padronização das instruções (GSP; n = 52) e com a padronização das instruções (GCP; n = 52).

No GSP, o aplicador entregava, recebia e conferia o preenchimento completo dos questionários, enquanto no GCP, o aplicador, antes de entregar o IFIS-BRA, fazia a leitura, em voz alta, das instruções padronizadas para todos os adolescentes. Caso houvesse alguma dúvida referente à compreensão, o terceiro aplicador direcionava-se ao participante e lia a definição do termo que foi previamente padronizada e especificada. Os grupos GSP e GCP, responderam o IFIS-BRA, duas vezes com um intervalo de duas semanas entre as aplicações^{5,11,14,15}.

Instrumentos

Antropometria

As avaliações antropométricas foram mensuradas conforme as recomendações da *International Society for the Advancement of Kinanthropometry*¹⁶⁻¹⁸. O Índice de Massa Corporal (IMC), foi determinado a partir da equação: $\text{Peso (kg)} / (\text{estatura (m)})^2$. Os resultados foram classificados conforme: baixo peso – 18.2; eutróficos – 18.6 a 24.9; sobrepeso – 25.0 a 29.0; obesidade grau I – 30.0 a 34.9; obesidade grau II – 35.0 a 39.9; e obesidade grau III – ≥ 40.0 ¹⁹.

Questionário de aptidão física

O questionário IFIS-BRA avalia a AFG e seus componentes – CC, FM, VA e Flex, respondidos em uma escala de resposta do tipo *Likert* de cinco pontos (“muito ruim”, “ruim”, “média”, “boa” e “muito boa”)^{5,20}. Na avaliação da confiabilidade, o mesmo apresentou um coeficiente de *Kappa* substancial (0.54 – 0.65)^{5,21}. No presente estudo, utilizamos o IFIS – BRA, acrescido da padronização das instruções. Essa versão apresentou evidências psicométricas adequadas ao conteúdo, clareza, compreensão e consistência interna¹².

Capacidade cardiorrespiratória

O teste 20m *shuttle run* (SR20m), avalia a capacidade cardiorrespiratória, sendo composto por 20 estágios, de 1 minuto. O primeiro estágio, inicia com velocidade de 8.5 km/h, sendo acrescidos 0.5 km/h, por meio

de um dispositivo sonoro, em cada estágio²². Baseado no último estágio completado, acrescentado da idade, o consumo máximo de oxigênio ($\text{VO}_{2\text{Máx}}$) foi predito²³. Na avaliação da confiabilidade (teste-reteste), o mesmo apresenta *Kappa* de 0,83²³.

Força muscular

Para avaliar a força de membros superiores, utilizamos o teste de força de preensão manual, avaliada por meio de um dinamômetro digital. Na avaliação da confiabilidade (teste-reteste), o mesmo apresenta *Kappa* de 0,77²³. O teste foi realizado apenas uma vez no braço direito (PMBD) e esquerdo (PMBE)⁵. A pontuação de força do membro superior foi calculada pela média da força de preensão das mãos direita e esquerda²³. Para avaliar a força de membros inferiores, utilizamos o teste de salto horizontal estático (SH)⁵. Na avaliação da confiabilidade (teste-reteste), o mesmo apresenta *Kappa* de 0.84²³. A força geral foi calculada pela média das pontuações da força nos testes PMBD, PMBE e SH²³.

Velocidade / agilidade

Para avaliar a velocidade/agilidade, foi utilizado o teste *shuttle run* 4x10m (SR10m)⁵. Na avaliação da confiabilidade (teste-reteste), o mesmo apresenta *Kappa* de 0.79²³. Foi permitido realizar duas tentativas com um intervalo de descanso de, no mínimo, dois minutos entre as tentativas, sendo registrado o melhor tempo^{5,23}.

Flexibilidade

Para avaliar a flexibilidade foi utilizado o teste de sentar e alcançar modificado (SA)^{23,24}. Na avaliação da confiabilidade (teste-reteste), o mesmo apresenta *Kappa* de 0.80²³.

Aptidão física geral

A aptidão física geral foi calculada como a média de todos os quatro componentes da aptidão física avaliada pelos testes físicos exibidos acima⁵. Na avaliação da confiabilidade (teste-reteste), o mesmo apresenta *Kappa* de 0.80²³.

Análise estatística

Validade de critério (objetivo 1)

Os dados foram inseridos por meio de digitação dupla, com checagem *a posteriori* no *software* Excel 2016, para *Windows* e analisados no SPSS 20.0 para *Windows*. Após a consolidação e validação dos dados, foi avaliada a normalidade pelo teste *kolmogorov-smirnov*, seguido das estatísticas descritiva e inferencial. Para as variáveis quantitativas foram utilizadas as medida de posição central (média), dispersão (desvio padrão) e precisão (intervalo de confiança 95%) e tamanho do efeito (*effect size*); para as variáveis qualitativas foi utilizada a distribuição de frequência.

Para comparação das variáveis entre os grupos, masculino e feminino, foi realizado o teste t de *student* independente. Para estimar o tamanho do efeito (*effect size*), foi aplicado o d *Cohen* seguido das classificações efeito pequeno ($d = 0.2$), efeito médio ($d = 0.5$) e efeito grande ($d = \geq 0.8$)²⁵. Com a finalidade de verificar evidências de validade de critério do tipo concorrente para o IFIS-BRA, inicialmente foi necessário ajustar as variáveis dos testes físicos, que foram utilizados como variáveis de critério (externas), para, em seguida, observar se houve correlação entre os testes físicos e os itens do IFIS – BRA autorrelatado. Para tanto, foram calculados os percentis dos testes físicos dos adolescentes participantes, considerando os percentis de adolescentes europeus em estudo anterior com a seguinte classificação em uma

escala do tipo *Likert* de cinco pontos: 1 = Muito ruim ($P < 20$); 2 = ruim ($P \leq 20$ a $P \leq 40$); 3 = Média ($P \leq 40$ a $P \leq 60$); 4 = Boa ($P \leq 60$ a $P \leq 80$) e 5 = Muito Boa ($P > 80$)⁵. Em seguida, foram calculadas análises de correlação de *Spearman*, recomendada para dados ordinais e estratificada por sexo e idade²⁶.

É importante ressaltar que se fazem necessários os ajustes para percentis, em virtude dos testes físicos, serem pré-estabelecidos e duplamente indiretos, tendo como resultados valores obtidos através de cálculos preditivos e em diversas unidades de medidas, explicando a necessidade das transformações e os principais resultados.

Confiabilidade (objetivo 2)

Os dados foram inseridos por meio de digitação dupla, com checagem *a posteriori* no *software* Excel 2016, para *Windows*; e analisados no *software* R versão 3.6.1 para *Windows*. A confiabilidade teste-reteste foi analisada por meio da média do percentual de concordância e pelo *kappa* ponderado quadrático. O percentual de concordância foi calculado pela diferença entre o teste inicial (T1), e o reteste (T2). Uma diferença (T2 - T1), igual a 0, foi classificada como concordância “perfeita” (mesma resposta teste-reteste), enquanto uma diferença de 0 ± 1 foi classificada como “perfeitamente aceitável”⁵. Para a classificação do *Kappa* ponderado quadrático foram utilizadas as seguintes classificações: < 0.00 (nenhum); 0.00 a 0.20 (leve); 0.21 a 0.40 (pequena); 0.41 a 0.60 (moderada); 0.61 a 0.80 (substancial) e 0.81 a 1.00 (quase perfeita)²¹.

Resultados

Validade de critério (objetivo 1)

A idade média foi de 15.00 ± 1.38 anos, sendo 59.80% ($n = 98$) do sexo feminino e 40.20% ($n = 66$), do sexo masculino. Referente à classificação do IMC de todos

os participantes, encontram-se: baixo peso (22.00%), eutróficos (57.30%), sobrepeso (12.20%), obesidade grau I (7.90%) e obesidade grau II (0.60%). Característica dos adolescentes participantes do estudo, dados antropométricos, na TABELA 1.

TABELA 1 - Características antropométricas dos adolescentes.

Variáveis	GG (n - 164)	GM (n - 66)	GF (n - 98)	Valor de p	Diferença do IC 95%	Tamanho do efeito
Idade (anos)	15.00 ± 1.38	14.86 ± 1.49	15.09 ± 1.32	0.30	-0.67 a 0.21	- 0,16
Estatutura (m)	1.63 ± 0.08	1.68 ± 0.07	1.59 ± 0.06	0.00*	0.07 a 0.11	1,38
MCT (kg)	58.65 ± 14.90	63.18 ± 17.37	55.9 ± 12.12	0.00*	3.04 a 12.14	0,50
IMC (kg / m²)	21.95 ± 4.56	22.16 ± 5.31	21.81 ± 4.00	0.65	-2.09 a 0.79	0,07
CC (cm)	70.60 ± 10.26	73.97 ± 12.10	68.33 ± 8.11	0.00*	2.53 a 8.75	-0,42

Os dados são apresentados em média e desvio padrão. As comparações entre os grupos (GG - grupo geral; GM - grupo masculino e GF - grupo feminino), são apresentadas pela diferença do intervalo de confiança 95% e tamanho do efeito. Valor de $p < 0,05^*$ diferença entre os grupos GM Vs GF (Teste t de Student independente). MCT - massa corporal total; IMC - índice de massa corporal total; CC - circunferência da cintura; m - metros; kg - quilograma; kg/m² - quilograma por metro ao quadrado; cm - centímetros.

Referente ao estágio de maturação sexual da Escala de *Tanner*, os itens de genitália do sexo masculino e feminino foram: nível um (0.00% e 20.40%), nível dois (4.50% e 22.40%), nível três (7.60% e 13.30%), nível quatro (65.20% e 31.60%) e nível cinco (22.70% e 12.20%). Os itens de maturação da mama do sexo feminino foram: nível dois (2.00%), nível três (20.40%), nível quatro

(39.80%) e nível cinco (37.80%).

Relativo às variáveis de aptidão física, observamos diferença entre os grupos, para as variáveis (SH, PMBD, PMBE, SR10m, SR20m, SR20mEST, VO_{2Máx}), com valores superiores no GM (TABELA 2). A mesma diferença foi observada nas variáveis de aptidão física autorrelatada (AFG, CC, FM e VA) (TABELA 3).

TABELA 2 - Comparação entre sexo dos testes de aptidão física de adolescentes.

Variáveis	GG (n - 164)	GM (n - 66)	GF (n - 98)	Valor de p	Diferença do IC95%	Tamanho do efeito
SH (cm)	147.65 ± 35.70	171.36 ± 34.14	131.67± 26.84	0.00*	30.26 a 49.12	1.29
PMBD (kgf)	28.74 ± 8.96	34.14 ± 10.41	25.11± 5.41	0.00*	6.57 a 11.49	1.08
PMBE (kgf)	27.45 ± 8.50	32.57 ± 10.00	24.01± 4.97	0.00*	6.63 a 10.89	1.08
SA (cm)	27.84 ± 8.36	26.40 ± 8.92	28.81± 7.86	0.07	-5.02 a 0.20	-0.28
SR10m (seg)	12.96 ± 2.32	11.62 ± 2.07	13.41± 2.21	0.00*	-2.47 a -1.11	-0.84
SR20m (seg)	138.59 ± 93.30	187.12 ± 106.17	105.90± 66.26	0.00*	54.63 a 107.81	0.92
SR20mEST (km/h)	9.33 ± 0.75	9.74 ± 0.84	9.06 ± 0.54	0.00*	0.47 a 0.89	0.96
VO _{2 Máx} (ml / kg / min)	34.07 ± 4.22	36.63 ± 4.12	32.35 ± 3.33	0.00*	3.13 a 5.43	1.14

Os dados são apresentados em média e desvio padrão. As comparações entre os grupos são apresentadas pela diferença do intervalo de confiança 95% e tamanho do efeito. Valor de $p < 0,05^*$ diferença entre os grupos GM Vs GF (Teste t de Student independente). SH - salto horizontal; PMBD - preensão manual do braço direito; PMBE - preensão manual do braço esquerdo; SA – sentar e alcançar; SR10m - shuttle run 4x10m; SR20m - 20m shuttle run test; SR20mEST - 20m shuttle run test por estágio; km/h - quilômetros por hora; kgf - quilograma-força; seg - segundos; ml/kg/min - mililitros de oxigênio por quilograma por minuto.

TABELA 3 - Análise de variáveis de aptidão física autorrelatada de adolescentes entre sexo.

Os dados são apresentados em média e desvio padrão. As comparações entre os grupos são apresentadas pela diferença do intervalo de confiança 95% e tamanho do efeito. Valor de $p < 0,05^*$ diferença entre os grupos GM Vs GF (Teste t de Student independente). AFG - aptidão física geral; CC - capacidade cardiorrespiratória; FM - força muscular; VA - velocidade/agilidade; Flex - flexibilidade.

Variáveis	GG (n - 164)	GM (n - 66)	GF (n - 98)	Valor de p	Diferença do IC95%	Tamanho do efeito
AGF	3.37 ± 0.79	3.53± 0.83	3.26 ± 0.75	0.03*	0.02 a 0.52	0.34
CC	3.23 ± 0.93	3.41± 0.96	3.11± 0.90	0.05*	0.01 a 0.59	0.32
FM	3.28 ±0.79	3.53± 0.68	3.11± 0.82	0.00*	0.16 a 0.68	0.55
VA	3.40 ± 0.87	3.61 ± 0.78	3.26 ± 0.90	0.01*	0.08 a 0.62	0.41
Flex	3.20 ± 0.84	3.23±0.87	3.17± 0.83	0.69	0.21 a 0.33	0.07

Para testar a correlação entre as respostas autorrelatadas, com a validade externa, foram realizados ajustes nos valores dos testes físicos, calculando os valores em percentis. Em seguida foram realizadas análises de correlação Rô de Spearman, para verificar validade de critério significativas entre todos os testes físicos e os itens do IFIS – BRA, especificamente observou-se correlação direta entre o teste de Força e o item Força do IFIS - BRA (Rô de Spearman 0.38; p 0.00);

correlação direta entre o teste de flexibilidade e o item flexibilidade do IFIS - BRA (Rô de Spearman 0.31; p 0.00); e, ainda, correlação direta entre o teste Cardiorrespiratório e o item Cardiorrespiratório do IFIS - BRA (Rô de Spearman 0.33; p 0.00). Finalmente, verificou-se correlação inversa entre o teste de Velocidade/Agilidade e o item Velocidade/Agilidade do IFIS - BRA (Rô de Spearman - 0.29; p 0.00), conforme o esperado, conforme (TABELA 4).

TABELA 4 - Resultados dos testes de aptidão física de acordo com as categorias do questionário IFIS - BRA.

Os dados são apresentados em média, desvio padrão, mínimo e máximo. n = 168. DP - desvio padrão; AFG - aptidão física geral; CC - capacidade cardiorrespiratória; FM - força muscular; VA - velocidade/agilidade; Flex - flexibilidade; U.A - unidade arbitrária; Min. - mínimo; Máx. - máximo.

Variáveis	Muito boa	Boa	Média	Ruim	Muito ruim
	Média e DP (Min. e Máx.)	Média e DP (Min. e Máx.)	Média e DP (Min. e Máx.)	Média e DP (Min. e Máx.)	Média e DP (Min. e Máx.)
CC (ml / kg / min)	40.70 ± 2.30 (37.23 a 44.56)	35.89 ± 0.59 (34.95 a 36.74)	33.73 ± 0.40 (33.30 a 34.30)	32.40 ± 0.77 (31.35 a 33.17)	29.07 ± 1.50 (25.53 a 30.32)
FM (U.A)	11.77 ± 12.98 (106.35 a 159.03)	97.06 ± 4.32 (91.70 a 105.78)	85.18 ± 2.79 (81.70 a 90.78)	77.08 ± 2.91 (72.33 a 81.33)	62.19 ± 9.87 (24.20 a 71.73)
VA (s)	16.32 ± 1.76 (15.00 a 20.00)	14.00 ± 0.00 (14.00 a 14.00)	13.00± 0.00 (13.00 a 13.00)	12.00 ± 0.00 (12.00 a 12.00)	10.46 ± 0.76 (7.00 a 11.00)
Flex (cm)	39.38 ± 3.39 (35.30 a 50.30)	32.49 ± 1.55 (30.30 a 35.00)	28.50 ± 1.01 (26.60 a 30.00)	23.92 ± 1.89 (21.00 a 26.30)	15.72 ± 3.75 (5.30 a 20.60)
AFG (U.A)	71.58 ± 6.54 (64.26 a 92.90)	60.01 ± 2.31 (57.13 a 63.86)	54.45 ± 1.27 (52.50 a 56.61)	50.12 ± 1.49 (47.44 a 52.29)	42.46 ± 5.03 (23.39 a 47.27)

Confiabilidade (objetivo 2)

A idade média, dos participantes de ambos os sexos, foi de 14.70 ± 1.50 anos, com a proporção de 51.90% (n = 94) do sexo masculino e 48.10% (n = 50), do sexo feminino. Para os grupos GCP e GSP foi

encontrada uma concordância média, de todos os itens, classificada como “perfeita” (64.20% e 58.50%), e “perfeitamente aceitável” (97.70% e 88.50%), respectivamente. Quando os itens foram avaliados separadamente, a concordância dos itens classificados como “perfeito” foram: AFG (63.50% e 61.50%);

CC (59.60% e 59.60%); FM (59.60% e 57.70%); VA (59.60% e 63.50%) e Flex (78.80% e 50.00%) e “perfeitamente aceitável”, foram: AFG (98.10% e 88.50%); CC (100.00% e 94.20%); FM (98.10% e

84.60%); VA (92.30% e 88.50%) e Flex (100.00% e 86.50%), respectivamente (FIGURA 1). A classificação e os resultados do *Kappa* ponderado quadrático do GCP foram superiores ao GSP (TABELA 5).

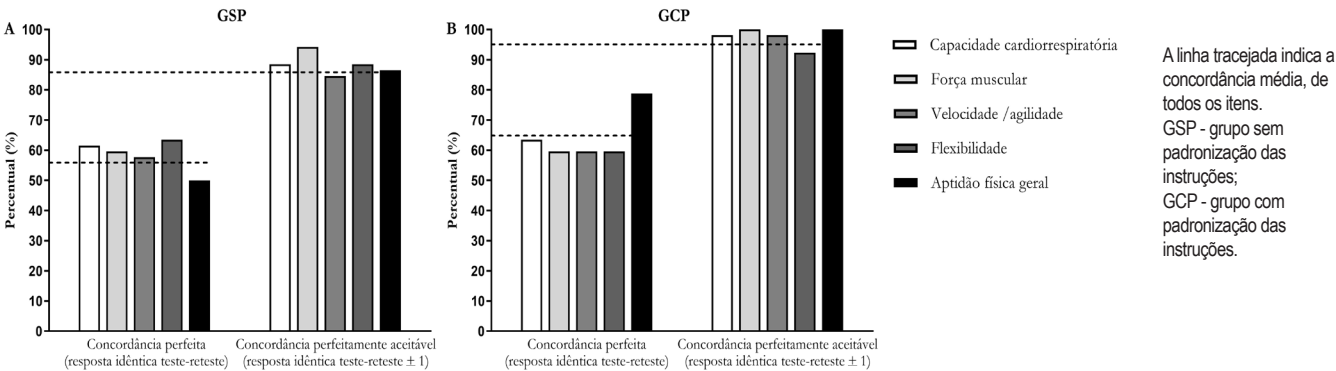


FIGURA 1 - Percentual de concordância entre o teste e o reteste.

TABELA 5 - Percentual de concordância entre o teste e o reteste.

Variáveis	GSP (n – 52)		GCP (n – 52)	
	Kappa ponderado	Classificação	Kappa ponderado	Classificação
AFG	0.50	Moderada	0.76	Substancial
CC	0.68	Substancial	0.80	Substancial
FM	0.37	Pequena	0.73	Substancial
VA	0.50	Moderada	0.63	Substancial
Flex	0.28	Pequena	0.89	Quase perfeita

Os dados são apresentados em média e *Kappa* ponderado quadrático. A classificação do *Kappa* ponderado quadrático adotada foi a de Landis JR e Koch GG. 1997²¹. AFG - aptidão física geral; CC - capacidade cardiorrespiratória; FM - força muscular; VA - velocidade/ agilidade; Flex - flexibilidade; GSP - grupo sem padronização das instruções; GCP - grupo com padronização das instruções.

Discussão

O presente estudo foi o primeiro a avaliar as evidências de validade baseada nas relações com medidas externas²⁷; e a confiabilidade do IFIS-BRA, com a padronização das instruções e forneceu evidências de sua aplicabilidade para classificar adolescentes de acordo com seus níveis de aptidão física. A partir dos resultados, podemos confirmar que os adolescentes compreenderam adequadamente a terminologia utilizada no IFIS - BRA, assistida pela padronização das

instruções; e foram capazes de relatar, com precisão, a aptidão física e seus domínios. De acordo com a literatura psicométrica, o coeficiente de validade de critério é calculado entre a medida (IFIS-BRA) e o critério (testes físicos), não havendo um parâmetro estabelecido para o tamanho que o coeficiente de validade deve assumir. Entretanto, pesquisadores indicam que este deve ser suficientemente grande para permitir que se identifique e se diferencie

o sujeito com relação à variável critério^{28,29}. Nesse sentido, correlações variando entre -0.29 e 0.37, apesar de não serem muito altas, devido à natureza das medidas indiretas, são consideradas suficientes para indicar que o IFIS - BRA possui evidência de validade de critério (evidências de validade baseada nas relações com medidas externas), uma vez que os resultados mostraram respostas condizentes entre a aptidão física autorrelatada dos adolescentes obtidas por meio do IFIS - BRA, quando relacionadas à aptidão física identificada ao realizar os testes físicos.

No estudo original, a validade baseada nas relações com medidas externas do IFIS³⁰, foi suficiente, sendo semelhante ao presente estudo. SANCHEZ-LOPEZ, et al.³¹, também obteve resultados semelhantes ao deste estudo, quando aplicado o IFIS em uma amostra próxima desta, crianças espanholas de 9 a 12 anos. Portanto, o presente estudo apresentou confiabilidade semelhante quando comparado com os estudos anteriores do IFIS.

Os resultados de aptidão física e seus domínios, mensurados pelos testes físicos e por instrumento autorrelatado (IFIS - BRA), são superiores no GM (SH, PMBD, PMBE, SA, SR10m, SR20m e SR20mEST, AFG, CC, FM e VA, exceto Flex e SA), que sugerem que são causados pelo estado maturacional dos adolescentes³².

Entre todos os estudos publicados com o IFIS, até o presente momento, este é o primeiro a estabelecer a padronização das instruções, para avaliar a confiabilidade do IFIS. Os resultados superiores do GCP sugerem que a padronização das instruções causa maior compreensão da terminologia usada nos itens e a ponderação dos adolescentes ao escutarem/lerem o cabeçalho e itens; visto que os adolescentes podem não associar a terminologia utilizada e a capacidade física requerida, mesmo que os itens referentes a cada aptidão física sejam autoexplicativos. A padronização das instruções em um questionário originalmente autorrelatado exige, da equipe, o treinamento prévio dos aplicadores e adição de tempo na coleta dos dados; porém, confere resultados mais confiáveis, conforme observamos em nosso estudo.

Para instrumentos que propõem avaliar relatos relacionados à saúde, é importante analisar sua confiabilidade; ou seja, a consistência de medição no tempo e no espaço, ou a partir de observadores diferentes, que podem ser obtidos por “teste-reteste”^{33,34}. Se um instrumento de avaliação tem baixa confiabilidade teste-reteste, significa que o mesmo não consegue distinguir mudanças reais nas pontuações, ao longo do tempo³⁴.

Além dessas considerações, estudos que utilizam questionários como ferramenta de avaliação, necessitam ter informações claras e concisas na escala *Likert* e nas respostas; visto que a incerteza de informações, contidas no questionário, pode induzir a respostas errôneas; assim, diminuindo a robustez dos resultados³⁵.

Nossos resultados de confiabilidade, a partir do *Kappa* ponderado quadrático, para todos os itens, foram superiores aos do estudo original, que também utilizou o IFIS em adolescentes⁵. Enquanto os valores deste estudo variaram de substancial à classificação quase perfeita de confiabilidade, os valores do estudo original variaram de moderada à classificação substancial, apresentando um *Kappa* ponderado quadrático por item: AGF (0.65), CC (0.58), FM (0.54), VA (0.60) e Flex (0.59)⁵. Apesar de apresentar níveis adequados de confiabilidade; possivelmente, se a padronização das instruções tivesse ocorrido no estudo preconizado por ORTEGA et al³⁰, seu nível de confiabilidade poderia ter sido mais alto⁵. PEREIRA¹¹ avaliou, por meio de Revisão Sistemática com Meta-análise, a confiabilidade teste-reteste (*Kappa* ponderado quadrático) do IFIS pelo subgrupo de adolescentes. A fidedignidade foi classificada como substancial: AGF (0.76), CC (0.80), FM (0.73), VA (0.76) e Flex (0.73). Portanto, o presente estudo apresentou confiabilidade superior/semelhante quando comparado com os estudos incluídos.

A avaliação da aptidão física de adolescentes, em larga escala, ainda é escassa, por motivos práticos, alto custo das ferramentas de avaliação, tempo elevado para a execução e necessidade da *expertise* do avaliador^{5,23}.

Assim, ferramentas não invasivas, indolores, de baixo custo e fácil aplicação, como os questionários, são meios alternativos; mas, é importante destacar que, antes de serem aplicados, os questionários devem passar por todas as etapas às quais o IFIS - BRA foi submetido, para verificar suas propriedades psicométricas a fim de garantir as evidências de validade dos seus resultados³⁵.

O IFIS-BRA com a padronização das instruções otimiza os resultados das medidas psicométricas de confiabilidade, assim assegurando a utilização, por profissionais da área da saúde e pesquisadores, como meio alternativo confiável para a avaliação da aptidão física de adolescentes brasileiros. Apesar do tamanho da padronização das instruções, faz-se necessário a aplicação desta, pois nem todos os adolescentes associam os itens do construto com capacidade física solicitada em virtude da falta de acesso às terminologias em seu cotidiano. O IFIS-BRA com a padronização das instruções atingiu bons níveis de desempenho psicométrico, assim assegurando a utilização, por profissionais da área da saúde e pesquisadores, como meio alterativo confiável para a avaliação da aptidão física de adolescentes brasileiros.

Apesar dos resultados dos itens do IFIS - BRA, com a padronização das instruções,

apresentarem aumento da confiabilidade no Brasil, este estudo ainda apresenta limitações, que devem ser detalhadas. A primeira limitação, como acontece com muitos questionários, é que todas as respostas são autorreferidas; significa que as respostas podem não corresponder, necessariamente, ao nível de aptidão física real dos participantes. No entanto, o autorrelato é um procedimento comum para coleta de dados, por meio de questionários e outros instrumentos; pois, o autorrelato é o método mais viável para coletar dados sobre percepções e crenças. A segunda limitação deste estudo consiste no fato de que a seleção dos participantes não foi aleatória. Essa falta de seleção aleatória pode significar que os participantes não eram verdadeiramente representativos da população em geral. Por fim, sugere-se que estudos em outros países deverão realizar a padronização da instrução, para analisar os efeitos, visto que a percepção de aptidão física pode ser modificada entre diferentes países.

Tomados em conjunto, podemos afirmar que a padronização da instrução do IFIS - BRA, oferece indicadores psicométricos satisfatórios de validade baseada nas relações com medidas externas e confiabilidade, sendo um instrumento simples e capaz de classificar corretamente a aptidão física de adolescentes brasileiros.

Agradecimentos

Agradecemos aos adolescentes que participaram do estudo, responsáveis dos adolescentes e gestores das escolas, que permitiram a realização deste estudo; e aos alunos que auxiliaram a coleta dos dados.

Conflitos de interesse

Declaramos que não há conflito de interesse.

Abstract

Evidence based on relations to other variables and reliability of the IFIS - BRA.

The assessment of the level of physical fitness in adolescents is a predictor of chronic degenerative diseases in adulthood, so its measurement by simple, inexpensive, and easily applicable means makes the International Fitness Scale of interest to public health worldwide. Therefore, our objective was to evaluate the evidence based on relations to other variables and reliability of the International Fitness Scale - Brazilian version (IFIS - BRA). Adolescents of both sexes participated in the study. The IFIS - BRA, and physical tests were applied to the adolescents. Taking the results together, we can state that the IFIS - BRA offers satisfactory psychometric indicators.

KEYWORDS: Physical fitness; Adolescent; Questionnaire; Validation study.

Referências

1. Schubert A, Januário RSB, Casonatto J, et al. Physical fitness and sports activities for children and adolescents. *Rev Bras Med Esporte*. 2016;22:142-146.
2. Brasil. Ministério da Saúde. Sistema de Informações sobre Mortalidade (SIM). 2021.
3. Lamounier JA, Abrantes MM. Prevalência de obesidade e sobrepeso na adolescência no Brasil. *Rev Médica Minas Gerais*. 2003;13:275-284.
4. Oliveira PCA. Testes físicos para avaliação da agilidade: possibilidade de adaptação ao futebol. *Rev Bras Futeb*. 2017;8:65-76.
5. Ortega FB, Ruiz JR, España-Romero V, et al. The International Fitness Scale (IFIS): usefulness of self-reported fitness in youth. *Int J Epidemiol*. 2011;40:701-711.
6. Pereira DA. Tradução, adaptação transcultural e validação do instrumento The International Fitness Scale para o uso no Brasil [dissertação]. Camaragibe (PE): Universidade de Pernambuco, Faculdade de Odontologia de Pernambuco; 2020.
7. Delmiro AS, Pinto NV. Evaluation of flexibility in jazz ballerines submitted to the proprioceptive neuromuscular facilitation method. *Res Soc Dev*. 2020;9(6):e146963568.
8. Silva JM, Neto AS, Santos GC, et al. Correlação entre a aptidão física, idade biológica e variáveis bioquímicas em adolescentes. *Rev Bras Prescrição Fisiol do Exerc*. 2019;13:1124-1130.
9. Hobold E, Junior LL, Campos FDS, et al. Comparison between indirect tests of cardiopulmonary fitness assessment and ergospirometry. *Cad Educ Física Esporte*. 2016;14:45-53.
10. Grant JS, Davis LL. Selection and Use of Content Experts for Instrument Development. *Res Nurs Health*. 1997;20:269-274.
11. Pereira DA, Correia JL, Carvas N, et al. Reliability of questionnaire The International Fitness Scale: a systematic review and meta-analysis. *Einstein*. 2020;18:1-9.
12. Pereira DA, Carvas JN, Correia JLL, et al. Transcultural adaptation: translation and validation of content of the brazilian version of the the international fitness scale instrument. *Adolescência Saúde*. 2019;16:80-92.
13. Pasquali L. Instrumentação Psicológica: Fundamentos e Práticas. 1. ed. Porto Alegre: Artmed; 2010.
14. Merellano-Navarro E, Collado-Mateo D, García-Rubio J, et al. Validity of the International Fitness Scale "IFIS" in older adults. *Exp Gerontol*. 2017;95:77-81.
15. Rufino OP, Garcia RJ, Merellano-Navarro E. Psychometric properties of the 'International Fitness Scale' in Chilean youth. *Retos*. 2016;31:23-27.
16. Marfell-Jones M, Stewart A, Olds T, et al. International Standards for Anthropometric Assessment. International Society for the Advancement of Kinanthropometry; 2011.
17. Lean MEJ, Han TS, Deurenberg P. Predicting body composition by densitometry from simple anthropometric measurements. *Am J Clin Nutr*. 1996;63:4-14.
18. Lohman TG, Roche AF, Martorell R. Anthropometric Standardization Reference Manual. Human Kinetics Books; 1988.

19. Silugem DM, Devinvenzi MU, Lessa AC. Diagnóstico do estado nutricional da criança e do adolescente. *J Pediatr*. 2000;76:275-284.
20. Pereira DA, Carvas Junior N, Correia Júnior JL, et al. IFIS - versão brasileira. 2021;1.
21. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977;33:159.
22. Leger LA, Mercier D, Gadoury C, et al. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci*. 1988;6:93-101.
23. Ramírez-Vélez R, Cruz-Salazar SM, Martínez M, et al. Construct validity and test-retest reliability of the International Fitness Scale (IFIS) in Colombian children and adolescents aged 9-17.9 years: the FUPRECOL study. *Peer J*. 2017;5:e3351.
24. Welk GJ, Meredith MD. *Fitnessgram/Activitygram reference guide*. 4. ed. The Cooper Institute; 2008.
25. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Elsevier. 2. ed. New York: Lawrence Erlbaum Associates; 1977.
26. Mukaka MM. A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012;24:69-71.
27. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. American Educational Research Association; 2014.
28. Cohen RJ, Swerdlik ME, Sturman ED. *Testagem e Avaliação Psicológica: Introdução a Testes e Medidas*. 8. ed. AMGH; 2014.
29. Pacico JC, Hutz CS. *Validade*. 1. ed. Psicometria: Artmed Editora; 2015. p. 188.
30. Ortega FB, Sánchez-López M, Solera-Martínez M, et al. Self-reported and measured cardiorespiratory fitness similarly predict cardiovascular disease risk in young adults. *Scand J Med Sci Sports*. 2013;23:749-757.
31. Sanchez-Lopez M, Martinez-Vizcaino V, Garcia-Hermoso A, et al. Construct validity and test-retest reliability of the International Fitness Scale (IFIS) in Spanish children aged 9-12 years. *Scand J Med Sci Sport*. 2015;25:543-551.
32. Vidal-Linhares R, Barros-Costa M, Fernandes FJ. The influence of sexual development on the basic physical characteristics of teenage boys. *Rev Salud Pública*. 2015;17:489-499.
33. Coluci MZO, Alexandre NMC, Milani D. Construção de instrumentos de medida na área da saúde. *Cien Saude Colet*. 2015;20:925-936.
34. Souza AC, Alexandre NMC, Guirardello EB. Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. *Epidemiol e Serviços Saúde*. 2017;26:649-659.
35. Keszei AP, Novak M, Streiner DL. Introduction to health measurement scales. *J Psychosom Res*. 2010;68:319-323.

ENDEREÇO

Ricardo de Freitas Dias
 Universidade de Pernambuco
 BR 203, Km 2, s/n - Vila Eduardo
 56.328-903 - Petrolina - PE - Brasil
 E-mail: ricardo.freitas@upe.br
 freitas-dias@hotmail.com

Submetido: 12/11/2022

Revisado: 05/01/2024

Aceito: 08/01/2024