

Cita: Pereira, Débora de Almeida; Correia Júnior, Jânio Luiz; Sobral, Hildeberto Dutra; Albuquerque da Costa, Julio Cezar; Gonçalves, Marina Pereira; Freitas -Dias, Ricardo (2023). Parámetros psicométricos do *International Fitness Scale* – versão brasileira via on-line e impresso. *Cuadernos de Psicología del Deporte*, 23(2), 26-37

Parámetros psicométricos de la escala *International Fitness Scale* – versión brasileña - online e impresso

Psychometric parameters of the *International Fitness Scale* – brazilian version - online and print

Parâmetros psicométricos do *International Fitness Scale* – versão brasileira via on-line e impresso

Pereira, Débora de Almeida¹, Correia Júnior, Jânio Luiz¹, Sobral, Hildeberto Dutra¹, Albuquerque da Costa, Julio Cezar², Gonçalves, Marina Pereira³, Freitas-Dias, Ricardo¹

¹ *Universidade de Pernambuco (UPE), Brasil;* ² *Universidade Federal de Alagoas (UFAL), Brasil;* ³ *Universidade Federal do Vale do São Francisco (Univasf), Brasil.*

RESUMEN

Con los avances tecnológicos, las encuestas en línea se han vuelto cada vez más comunes, lo que hace viable la evaluación del estado físico en línea. Nuestro objetivo fue evaluar las pruebas con base en la estructura interna, verificar la invariabilidad factorial del modelo entre el formato online y la versión impresa (papel y bolígrafo), así como probar la fiabilidad de la *International Fitness Scale* - versión brasileña (IFIS - BRA). Participaron en el estudio adolescentes (n = 294), con edades comprendidas entre 12 y 17 años, de ambos sexos. Los adolescentes respondieron la versión impresa y en línea del IFIS - BRA en dos momentos, con un intervalo de dos semanas entre aplicaciones. La confiabilidad de todos los elementos se clasificó como "sustancial". Se encontró una concordancia media para todos los ítems, clasificados como "perfecto" (69,7%) y "perfectamente aceptable" (97,1%). Podemos afirmar que la versión online del IFIS - BRA, alcanzó niveles adecuados de desempeño psicométrico, asegurando así el uso, por parte de profesionales del área de la salud e investigadores, como medio alterno confiable para la evaluación de la aptitud física de los adolescentes brasileños.

Palabras clave: Aptitud Física; Adolescente; Cuestionario; Autoinforme; Sistemas On-Line.

ABSTRACT

With technological advances, online surveys have become increasingly common, which makes online fitness assessment feasible. We aimed to evaluate the evidence based on the internal structure, to verify the factorial invariance of the model between the online format and the printed version (paper and pen), as well as to test the reliability of the *International Fitness Scale* - Brazilian version (IFIS-BRA). Participants were adolescents (n = 294), aged between 12 to 17 years, of both sexes. Adolescents answered the printed and online version of the IFIS - BRA in two moments, with an interval of two weeks between applications. Reliability of all items was classified as "substantial". A mean agreement was found for all items, classified as "perfect" (69.7%) and "perfectly acceptable" (97.1%). We can affirm that the online version of the IFIS - BRA, reached adequate levels of psychometric performance, thus assuring its use by health professionals and researchers as a reliable alternative means of assessing physical fitness in Brazilian adolescents.

Keywords: Physical Fitness; Adolescent; Questionnaire; Self Report; Online Systems.

Parâmetros psicométricos do IFIS – BRA

RESUMO

Com os avanços tecnológicos as pesquisas *on-line* tornaram-se cada vez mais comuns, o que torna a avaliação da aptidão física *on-line* viável. Objetivamos avaliar a evidência baseada na estrutura interna, verificar a invariância fatorial do modelo entre o formato *on-line* e à versão impressa (papel e caneta), bem como testar a confiabilidade do *International Fitness Scale* – versão brasileira (IFIS-BRA). Participaram do estudo adolescentes (n = 294), na faixa etária entre 12 e 17 anos, de ambos os sexos. Os adolescentes responderam à versão impressa e *on-line* do IFIS – BRA em dois momentos, com o intervalo de duas semanas entre as aplicações. A confiabilidade, de todos os itens, foi classificada como “substancial”. Foi encontrada uma concordância média, de todos os itens, classificada como “perfeita” (69,7%), e “perfeitamente aceitável” (97,1%). Podemos afirmar que a versão *on-line* do IFIS – BRA, atingiu níveis adequados de desempenho psicométrico, 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.

Palavras-chave: Aptidão física; Adolescente; Questionário; Autorrelato; Sistemas Online.

INTRODUÇÃO

Pesquisas sobre a avaliação da aptidão física são tradicionalmente realizadas por testes físicos; porém, seu uso é limitado, pois necessita de especialistas, grande intervalo de tempo para execução e apresenta dificuldade na avaliação de grandes grupos. Assim, uma alternativa está focada nos questionários impressos (Ortega et al., 2011). Com os avanços tecnológicos, as pesquisas *on-line* tornaram-se cada vez mais comuns; desde o recrutamento dos participantes pelas mídias sociais até a aplicação de questionários autorrelatados, via *on-line*; o que torna a avaliação da aptidão física em larga escala, por meio digital, uma alternativa viável (Belisario et al., 2015; Hirao et al., 2021; Zeleke et al., 2019).

A aplicação de questionários *on-line* apresenta um formato mais atrativo e possui inúmeras vantagens, em comparação aos métodos tradicionais de coleta de dados em papel e caneta (versão impressa). Dentre elas, podemos destacar: o custo reduzido, capacidade ilimitada de armazenamento, melhor gerenciamento e distribuição dos dados, maior velocidade na obtenção de dados, o registro mais preciso das respostas, garantia do preenchimento completo do questionário, maior segurança no armazenamento das informações coletadas, acesso mais fácil aos participantes – especialmente em áreas de difícil acesso geográfico – maior sensação de anonimato e menor impacto ecológico (Belisario et al., 2015; González-Ruiz et al., 2010, 2015; Reigal et al., 2020; Zeleke et al., 2019).

Apesar do grande potencial do uso da *internet* como ferramenta para coletar informações sobre a aptidão física (Bicen et al., 2020), atualmente não existe um

consenso na literatura sobre a validade/confiabilidade desse tipo de questionário *on-line* (Belisario et al., 2015; Stawarz et al., 2018). Estudos que revisaram sistematicamente a coleta de dados, pela *internet*, geralmente focaram na confiabilidade (Braekman et al., 2018; Campos et al., 2011) e na taxa de resposta (Ebert et al., 2018). A maioria desses estudos descreve pequenas diferenças entre os formatos utilizados para administrar o questionário.

A literatura científica apresenta que há equivalência, e índices satisfatórios de estimativas de confiabilidade e evidência baseada na estrutura interna entre instrumentos aplicados via papel e caneta e *on-line* (Donovan et al., 2000; Joubert & Kriek, 2009; King & Miles, 1995; Mead & Drasgow, 1993); porém, alguns instrumentos podem sofrer uma diminuição significativa da validade de construto, da evidência baseada na estrutura interna e da confiabilidade quando adaptados à *internet*, pois os resultados podem sofrer alterações a depender da facilidade, ou dificuldade, com que os respondentes interagem e compreendem a forma de apresentação do questionário *on-line* (Buchanan et al., 2005; Finegan & Allen, 1994; Gelder et al., 2010). Portanto, apesar dos inúmeros benefícios supracitados da administração dos questionários *on-line*, os níveis de validade, de construto e confiabilidade, são variados e inconsistentes, o que pode, conseqüentemente, repercutir na validade interna e externa do instrumento e da pesquisa.

Até o presente momento, não há estudo publicado sobre a validade de construto e confiabilidade da administração *on-line*, do questionário *International Fitness Scale* (IFIS), que avalia, em adolescentes, a aptidão física geral (AFG) e seus domínios, capacidade cardiorrespiratória (CC), força muscular

(FM), velocidade/agilidade (VA) e flexibilidade (Flex) (Ortega et al., 2011).

Pereira et al., 2019, recentemente, adaptaram e apresentaram evidências de validade de conteúdo do IFIS para o idioma português brasileiro (Pereira et al., 2019). Este é um questionário que apresenta boas propriedades psicométricas, de fácil aplicabilidade e com baixo custo, podendo ser aplicado em larga escala geográfica e por qualquer profissional de saúde (Pereira et al., 2020).

Considerando que o sedentarismo e a baixa aptidão física, na adolescência, são fatores que influenciam no surgimento de doenças crônico-degenerativas (diabetes *mellitus*, dislipidemias, aterosclerose, hipertensão arterial e alguns tipos de câncer) (Ortega et al., 2011; Ortega & Montero, 2021; Zhang et al., 2020), na vida adulta, torna-se oportuno, e de interesse da saúde pública mundial, propor ferramentas que avaliam a aptidão física em larga escala (American College of Sports Medicine, 2017; Falconi et al., 2019; World Health Organization, 2020). Devido à necessidade identificada na literatura, o objetivo do estudo foi avaliar a evidência baseada na estrutura interna, verificar a invariância fatorial do modelo entre o formato *on-line* e à versão impressa (papel e caneta), bem como testar a confiabilidade do *International Fitness Scale* – versão brasileira (IFIS-BRA).

MÉTODOS

Participantes

Neste estudo, foram selecionados por conveniência 294 adolescentes, com a faixa etária entre 12 e 17 anos de ambos os sexos. Os adolescentes que atenderam aos critérios de inclusão foram convidados, por pesquisadores treinados, para participar do estudo, por meio de convites verbais. Foram determinados, como critérios de inclusão, adolescentes na faixa etária entre 12 e 17 anos. Foram excluídos os voluntários que não realizaram o preenchimento completo do questionário IFIS – BRA; os que não compareceram à coleta de dados, no reteste (02 semanas após a primeira aplicação); e os voluntários que apresentaram comprometimento das condições de saúde física ou intelectual, declarada, previamente, pelo professor e adolescentes gestantes.

Tamanho da amostra

Conforme proposto anteriormente, para a validação de instrumentos psicométricos, recomenda-se pelo menos 10 indivíduos por fator avaliado, ou 20 a 30 indivíduos por itens ou variáveis, ou seja, 50 a 150 indivíduos conforme o questionário IFS – BRA (Marôco, 2010).

Instrumento

Para avaliar os adolescentes, utilizamos o questionário *The International Fitness Scale* – versão brasileira (IFIS – BRA). Este instrumento avalia a aptidão física geral e seus componentes (CC, FM, VA e Flex), utilizando uma escala de resposta de cinco pontos, que varia de “muito ruim” a “muito boa”. No presente estudo, utilizamos a versão validada do IFIS, para o idioma português do Brasil, por Pereira et al., 2019 (Pereira et al., 2019).

Procedimentos

Este projeto foi aprovado pelo Comitê de Ética em Pesquisa, com o objetivo de respeitar o disposto na Resolução nº 466/12 do Conselho Nacional de Saúde. 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.

Para a aplicação do questionário impresso (papel e caneta), utilizamos o IFIS – BRA (Pereira et al., 2021). Enquanto para a aplicação do IFIS – BRA de forma *on-line*, os componentes (cabeçalho, domínios e a escala) do questionário foram inseridos no *Google Forms*, por um pesquisador (R.F.D.), com experiência na plataforma. O *Google Forms* é um serviço gratuito para criar formulários *on-line*, que está incluído no pacote do *Google Drive*.

Para garantir que todas as respostas fossem respondidas e enviadas, todas as perguntas na versão *on-line* eram obrigatórias. Após o envio do questionário, os dados foram carregados automaticamente, em um banco de dados, no qual cada coluna correspondia a um item e cada linha correspondia a um participante.

Por fim, para garantir a reprodutibilidade das informações, que seriam passadas aos participantes,

Parâmetros psicométricos do IFIS – BRA

na aplicação do IFIS – BRA (Pereira et al., 2019), todos os aplicadores passaram por um treinamento de calibração, no qual receberam as instruções, sobre a forma de aplicação.

Os adolescentes responderam ao questionário em dois momentos independentes: T1 – aplicação da versão impressa (papel e caneta) e T2 – aplicação *on-line*, no qual o período de intervalo entre a aplicação de T1 e T2 foi de duas semanas (Figura 1). Optamos por este período, por ser o intervalo de tempo que apresentou os melhores resultados de reprodutibilidade, em amostras com características semelhantes às deste estudo (Pereira et al., 2020), além de ser o período recomendado na literatura científica (Polit, 2014).

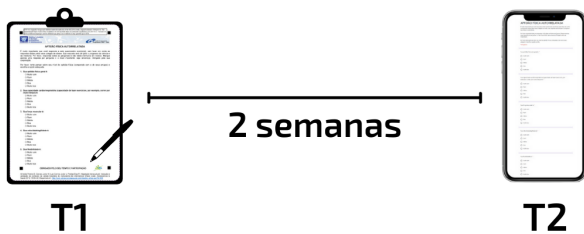


Figura 1. Aplicação do questionário no formato presencial e *on-line*. Legenda: T1 – teste (aplicação 1); T2 – reteste (aplicação 2).

Para garantir que as respostas não seriam copiadas entre os adolescentes, para a aplicação do questionário, os adolescentes foram dispostos em carteiras com divisórias, de modo que não houvesse interação entre os mesmos. Dois aplicadores eram responsáveis por entregar, receber, conferir o preenchimento completo dos questionários e fiscalizar a interação entre os adolescentes. O questionário no momento T1 e T2 foi aplicado em apenas 10 adolescentes por vez.

Para a aplicação do questionário *on-line*, foi disponibilizado um dispositivo móvel (*smartphone* ou *tablet*), com o *link* do IFIS – BRA *on-line*, no *google forms*. Vale ressaltar que não houve interferência dos aplicadores, na coleta *on-line*; e, após o término da aplicação, foi conferido o envio do questionário *on-line*. Para garantir a reprodutibilidade da coleta, as funções dos aplicadores e o local foram os mesmos durante as aplicações (T1 e T2).

Análise dos dados

Os dados foram inseridos por meio de digitação dupla, com checagem *a posteriori*, no *software* Excel 2016, para *Windows*; e analisados no pacote *Lavaan* (Rosseel, 2014) e *sem Tools* (Jorgensen et al., 2021), do programa do *software* R, para *Windows*. Após validação e consolidação dos dados, foi executado teste *Kolmogorov-Smirnov*, para avaliar a normalidade destes dados, seguido da análise descritiva e inferencial. As variáveis foram sintetizadas em medidas de tendência central (média), dispersão (desvio padrão) e posição (intervalo de confiança 95% e *effect size*). Para comparação intra-grupos (IFIS – BRA *on-line* Vs IFIS –BRA versão impressa), das variáveis, foi realizado o teste T de *student* dependente, com nível de significância estabelecido de 5% ($p < 0,05$). Para estimar o *effect size*, foi utilizado o *r* de *Pearson*, com as classificações: efeito pequeno ($r = 0,10$), efeito médio ($r = 0,30$), efeito grande ($r = \geq 0,50$) (Cohen, 1977). Em seguida, a validade do construto foi analisada por meio da análise fatorial confirmatória (AFC), utilizando o estimador *Weighted Least Squares Mean and Variance-Adjusted* (WLSMV) por se tratar de uma matriz policóricas dos dados (Li, 2016; Muthén & Muthén, 2014). A qualidade do modelo foi testada utilizando índices de ajuste: a proporção de χ^2/df , comparativa índice de ajuste (CFI), erro quadrático médio aproximado da raiz (RMSEA) e seu respectivo intervalo de confiança de 95% (RMSEA 95%). A invariância do instrumento foi mensurada a partir da Análise Fatorial Confirmatória Multigrupo (AFCMG).

A proporção de $\chi^2/$ grau de liberdade é um teste de qualidade do ajuste. Valores acima de 5 correspondem a um ajuste inadequado; valores entre 2 e 5, correspondem a um ajuste aceitável; valores entre 1 e 2 correspondem a um ajuste próximo; e valores abaixo de 1, correspondem a um bom ajuste (Marôco, 2014).

O CFI avalia a adaptação do modelo teórico, em comparação com o modelo nulo, quando este modelo nulo é independente da dimensão da amostra. Valores de $CFI \leq 0,8$ indicam um ajuste inadequado; valores entre 0,8 e 0,9, indicam um ajuste aceitável; valores entre 0,90 e 0,95 indicam um bom ajuste; e valores $\geq 0,95$, indicam um ajuste excelente (Marôco, 2014).

O RMSEA avalia se o modelo ajustado está aproximadamente correto em comparação com o ajuste obtido; se o mínimo da função de discrepância

foi obtido a partir de valores populacionais. Os valores RMSEA (IC95%: p-valor: $RMSEA \leq 0,05$), acima de 1,0 sugerem um ajuste inaceitável; valores entre 0,05 e 0,10 sugerem um ajuste aceitável; e valores de $\leq 0,05$ sugerem um excelente ajuste (Marôco, 2014). A variância média extraída foi calculada para avaliar a proporção de variação dos itens que são explicados pelo fator ao qual eles pertencem. Uma convergência positiva do modelo é assumida, se os valores de variação média, extraídos, forem iguais ou acima de 0,5.

O *Tucker-Lewis coeficiente* (TLI), é considerado um indicador global de adequação ao modelo, no qual, valores acima de 0,90 estão de acordo com o recomendado pela literatura (Hair et al., 2009). O *Standardized Root Mean Square Residual* (SRMR) é uma média dos resíduos observados e estimados, de covariância da variância; um índice de ajuste absoluto, no qual, valores abaixo 0,08 estão de acordo com o recomendado pela literatura (Brown, 2015; Tabachnick & Fidell, 2012).

Para a Análise Fatorial Confirmatória, quanto menor o valor do χ^2 (e maior o valor de significância obtido a partir da estatística $p < 0,05$), melhores são os resultados; pois esse achado indica que não há nenhuma diferença estatisticamente significativa entre as matrizes observada e esperada (Hair et al., 2009).

Ademais, a Análise Fatorial Confirmatória Multigrupo (AFCMG) foi investigada a partir dos pontos de corte: $\Delta CFI < 0,01$ e $\Delta RMSEA < 0,015$ (Wu et al., 2007). Os tipos de invariância fatorial aqui investigados foram: a) Invariância configural: checka se a estrutura fatorial do instrumento se mantém nos dois grupos estudados; b) Invariância métrica: analisa se as cargas fatoriais dos dois grupos são estatisticamente equivalentes; c) Invariância escalar: infere que os escores estão relacionados ao nível de traço latente em todos os grupos; d) Invariância residual: avalia se os resíduos dos itens são estatisticamente iguais entre os grupos (Damásio, 2013; Putnick & Bornstein, 2016).

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”.

Para a classificação do *Kappa* ponderado quadrático, foram utilizadas as seguintes classificações: $< 0,00$ (nenhum); 0,00 – 0,20 (leve); 0,21 – 0,40 (pequena); 0,41 – 0,60 (moderada); 0,61 – 0,80 (substancial) e 0,81 – 1,00 (quase perfeita) (Landis & Koch, 1977). Por fim, foi verificada a confiabilidade das versões presencial e *on-line*, por meio de análise de consistência interna (Alfa de Cronbach e Ômega), sendo considerados como adequados valores iguais ou superiores a 0,70 (Tabachnick & Fidell, 2012), seguindo as classificações: $< 0,51$ (inaceitável); 0,51 – 0,61 (pobre); 0,61 – 0,71 (questionável); 0,71 – 0,81 (aceitável); 0,81 – 0,90 (bom) e $> 0,91$ (excelente) (George & Mallery, 2003).

RESULTADOS

A amostra foi composta por 294 adolescentes de ambos os sexos, com idade média de $14,7 \pm 1,5$ anos; sendo: 46,2% do sexo masculino, e 53,8% do sexo feminino. A pontuação obtida nos itens do

questionário aplicado via *on-line* e impresso, foram semelhantes para todos os itens, exceto para a Aptidão física geral (Tabela 1).

Posteriormente, a Análise Fatorial Confirmatória (AFC) sustentou a hipóteses do fator geral composto pelos cinco itens do IFIS-BRA, quando investigada a partir da amostra de respondentes com o questionário impresso e *On-line* [$\chi^2(5) = 4,112$ ($p = 0,533$), $\chi^2/gl = 0,822$, CFI = 0,985, TLI = 0,969, SRMR = 0,024 e RMSEA = 0,054 (IC90% = 0,020 - 0,089)].

Ressaltamos, ainda, que todos os itens do questionário apresentaram saturações (lambdas), estatisticamente diferente de zero ($\lambda \neq 0$; $p < 0,001$), com valores variando de 0,451 (item 5) a 0,836 (item 1) para a versão impressa (Figura 2A) e valores variando de 0,496 (item 5) a 0,781 (item 1) para a versão *on-line* (Figura 2B); indicando estabilidade de estimativas fatoriais e ajuste do modelo para os dados, o que permite supor que ambas as versões (impressa e *on-line*) do IFIS-BRA apresentam validade baseada na estrutura interna.

Parâmetros psicométricos do IFIS – BRA

Tabela 1. Resultados do IFIS – BRA aplicado em adolescente nas versões impressa e *on-line* (n=294).

Itens	Versão <i>on-line</i>	Versão impressa	Valor de p	IC95%	Effectsize
Aptidão física geral	3,50 ± 0,86 (1 e 5)	3,41 ± 0,86 (1 e 5)	<0,05*	-0,05 a 0,23	0,05
Cardiorrespiratória	3,25 ± 0,94 (1 e 5)	3,20 ± 0,96 (1 e 5)	>0,05	-0,10 a 0,20	0,03
Força muscular	3,31 ± 0,84 (1 e 5)	3,32 ± 0,89 (1 e 5)	>0,05	-0,15 a 0,13	0,00
Velocidade/agilidade	3,55 ± 0,85 (1 e 5)	3,59 ± 0,88 (1 e 5)	>0,05	-0,18 a 0,10	0,02
Flexibilidade	3,20 ± 0,94 (1 e 5)	3,21 ± 0,95 (1 e 5)	>0,05	-0,16 a 0,14	0,00

Os dados são apresentados em média, desvio padrão, máximo e mínimo, intervalo de confiança 95% e *effect size*. *p < 0,05 – teste T *student* dependente.

Entretanto, optou-se por investigar a invariância fatorial das duas versões do IFIS-BRA por meio da Análise Fatorial Confirmatória Multigrupo (AFCMG), que atestou a equidade do instrumento entre a versão Online e sua versão Impressa, a partir da invariância configural, que infere que a estrutura fatorial se mantém a mesma nos dois tipos de aplicação (Tabela 2).

No que se refere aos índices de consistência interna, os valores de alfa de *Cronbach* e Ômega de *McDonald* foram adequados ($\alpha = 0,78$; $\omega = 0,79$). Nossos achados sugerem que ambas as formas apresentam índices de consistência interna aceitáveis. Para as versões impressa e *on-line*, foi encontrada uma concordância média de todos os itens, classificada como “perfeita” (69,7%); e “perfeitamente aceitável” (97,1%), respectivamente. Quando foram avaliados separadamente, a concordância dos itens classificados como “perfeito” foram: AFG (71,1%); CC (66,3%); FM (67,3%); VA (74,5 %) e Flex (69,0%). Os itens classificados como “perfeitamente aceitável”, foram: AFG (95,9%); CC (97,6%); FM (97,3%); VA (98,3%) e Flex (96,6%), respectivamente (Figura 3). Os resultados do *Kappa* ponderado quadrático foram classificados como “substancial” (Tabela 3).

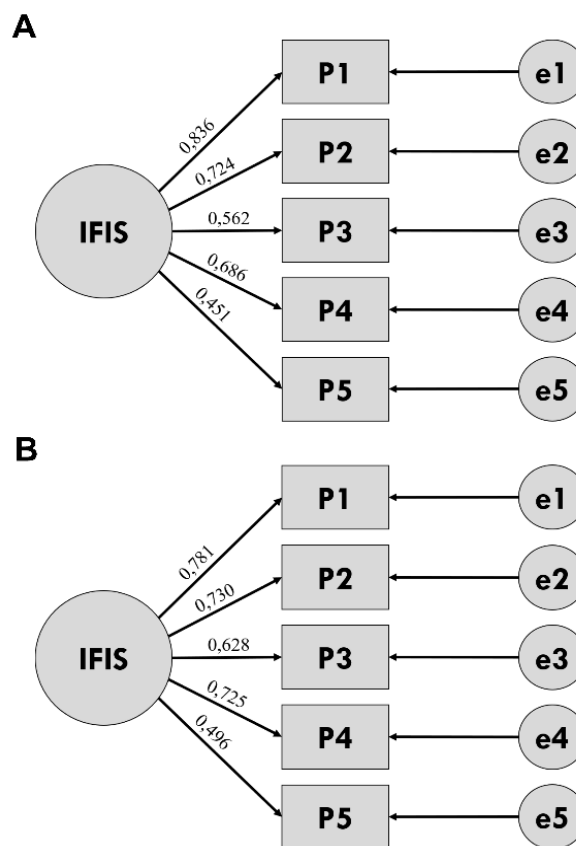


Figura 2. Estrutura Fatorial e Cargas Fatoriais do IFIS-BRA. A) Modalidade impressa. B) Modalidade *on-line*.

Tabela 3. Confiabilidade do IFIS – BRA, versões impressa e *on-line*.

Desfecho	<i>Kappa</i> ponderado quadrático	IC95%	Classificação
Aptidão física geral	0,71	0,64 a 0,78	Substancial
Cardiorrespiratória	0,76	0,70 a 0,82	Substancial
Força muscular	0,73	0,66 a 0,79	Substancial
Velocidade/agilidade	0,79	0,74 a 0,85	Substancial
Flexibilidade	0,74	0,67 a 0,82	Substancial

Os dados são apresentados no valor de *Kappa* ponderado quadrático e o intervalo de confiança, a 95%. A classificação do *Kappa* ponderado quadrático adotada foi a de Landis JR e Koch GG (1997) (Landis & Koch, 1977).

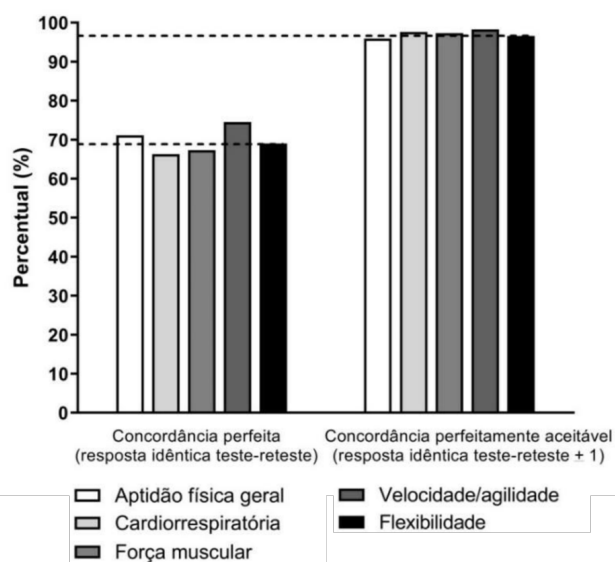


Figura 3. Percentual de concordância entre as versões impressa e *on-line*.

DISCUSSÃO

O presente estudo foi o primeiro a avaliar a validade baseada na estrutura interna (validade fatorial) e consistência interna do *International Fitness Scale* (IFIS – BRA) e forneceu evidências de sua aplicabilidade no contexto *on-line*.

O modelo de único fator composto pelos cinco itens do IFIS – BRA, tanto na versão *on-line*, quanto na versão impressa, mostra consistência interna aceitável no total de itens, com valores de alfa de *Cronbach* e coeficiente Ômega de *McDonald*,

adequados para as versões (impressa e *on-line*); sendo que nossos resultados corroboraram com estudos anteriores que, também, identificaram valores de alfa no IFIS considerados adequados (acima de 0,70) (Espanol-Moya & Ramirez-Vélez, 2014; Pereira et al., 2019; Ramirez-Vélez et al., 2017).

Ressaltamos que este método tem por objetivo indicar a avaliação da consistência interna, por meio da magnitude de covariância entre os itens, ou seja, quanto maior for o valor do alfa, os itens, individualmente, produzem baixa variabilidade de respostas, caracterizando-se como itens homogêneos (Tabachnick & Fidell, 2012).

Apesar dos resultados da estatística χ^2 do IFIS – BRA, versão *on-line*, ser significativa, explicitando uma diferença entre as matrizes, podemos atestar bons ajustes para as duas versões (impressa e *on-line*), conforme apontam os resultados de natureza incremental (TLI e CFI), e de natureza absoluta (SRMR e RMSEA), em que todos os índices alcançaram resultados adequados, conforme a literatura.

Assim, os achados deste estudo são semelhantes aos encontrados em outras versões do IFIS, que foram traduzidas para outros idiomas (Ortega et al., 2011). Em relação à estrutura fatorial das versões impressa e *on-line* do IFIS – BRA, ambas apresentaram índices de ajustes satisfatórios, os quais confirmam a estrutura unifatorial e validade de construto deste instrumento, corroborando com a versão original do IFIS, nas demais versões linguísticas (Ortega et al., 2011).

Vários estudos têm mostrado que a aplicação de questionários e outros instrumentos de coleta de dados *on-line* não comprometem suas medidas

Parâmetros psicométricos do IFIS – BRA

(Braekman et al., 2018; Campos et al., 2011; Hedman et al., 2010). Por outro lado, relatos de resultados mais positivos dos instrumentos *on-line*, quando comparados aos tradicionais impressos, também são descritos na literatura (Campos et al., 2011).

A hipótese é de que a aplicação de questionários, pela *internet*, reduza a tendência de fornecer respostas socialmente desejáveis. Isso porque os computadores podem proporcionar uma situação mais impessoal, em que os participantes se sintam mais anônimos e menos preocupados em como podem ser percebidos por outras pessoas (Campos et al., 2011). Além destas considerações, observamos que a ausência de um aplicador para fornecer instruções mais assertivas de preenchimento dos questionários podem comprometer as respostas; por isso, a necessidade de verificar evidências de validade e fidedignidade tanto da versão impressa, quanto da versão *on-line* (Álvarez-Gallardo et al., 2016; Ramírez-Vélez et al., 2017).

Uma das principais formas de se aferir a confiabilidade de um instrumento consiste na aplicação de uma mesma medida, em dois momentos distintos, de modo a encontrar resultados semelhantes. Assim, a forma mais indicada na literatura, para comprovação, consiste no método de teste-reteste. Este método tem como objetivo avaliar a equivalência das respostas dos itens de um instrumento, ou seja, significa que o mesmo teste, medindo os mesmos sujeitos, em ocasiões diferentes, em condições perfeitas, produzirá resultados idênticos. Assim, instrumentos com baixa confiabilidade teste-reteste, não conseguem evidenciar mudanças ao longo do tempo (Pereira et al., 2020; Souza et al., 2017).

Em nosso estudo, não encontramos diferenças entre os resultados das respostas dos participantes, nos dois formatos impresso e *on-line* do IFIS – BRA (Tabela 1). Encontramos uma concordância média, de todos os itens, classificada como “perfeita” e “perfeitamente aceitável” (Figura 3).

Com relação à confiabilidade, este estudo apresentou uma classificação substancial, em todos os domínios, com um *Kappa* ponderado quadrático substancial (Tabela 2), e superior ao original de Ortega et al. (2011) (Ortega et al., 2011), que aplicou o IFIS presencialmente e autorrelatado. Pereira et al. (2020) (Pereira et al., 2020), verificaram, na revisão sistemática e meta-análise, onde avaliaram a confiabilidade do IFIS, que os resultados na análise,

por subgrupo de adolescentes, foram classificados como “substancial”: AGF (0,76), CC (0,80), FM (0,73), VA (0,76) e Flex (0,73).

Em suma, quando comparado com estudos apresentados na análise por subgrupo de adolescentes, o presente estudo apresentou maior confiabilidade nos itens velocidade/agilidade e flexibilidade, menor confiabilidade nos itens aptidão física geral e capacidade cardiorrespiratória e confiabilidade igual no item força muscular (Pereira et al., 2020), indicando excelente estabilidade nas respostas, obtidas por meio do formato *on-line* do IFIS–BRA, em comparação ao tradicional formato impresso (papel e caneta).

A avaliação do nível de aptidão física de adolescentes, em larga escala, ainda é um desafio, pelos motivos associados às características de questionários aplicados presencialmente. (Chirivella, 2016; Lopes & Sá, 2021; Pans et al., 2021) Assim, ferramentas, como os questionários aplicados via *on-line*, são fundamentais; mas, é importante destacar que, antes de serem aplicados em diferentes condições, os questionários devem apresentar diferentes tipos de evidências de validade, para seu uso, no contexto pretendido, tais como: a validade de conteúdo, construto e critério, para uso no Brasil (Pereira et al., 2019), a fim de garantir a validade e confiabilidade dos seus resultados.

Limitações e expectativas futuras

Apesar dos resultados, com a aplicação via *on-line*, do IFIS – BRA serem promissores, 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. Outra 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 a utilização das versões *on-line* e impressa da IFIS-BRA a fim de averiguar a

validade convergente do instrumento em suas duas versões de aplicação, para se obter indicadores de equivalência entre os contextos de aplicação.

Implicações práticas

O IFIS – BRA via *on-line* atingiu bons níveis de desempenho psicométrico das medidas 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.

CONSIDERAÇÕES FINAIS

Tomados em conjunto, os resultados, podemos afirmar que a versão *on-line* do *International Fitness Scale* – versão brasileira (IFIS – BRA), atingiu bons níveis de desempenho psicométrico. A solução fatorial, encontrada por meio do AFC, foi semelhante às apresentadas no formato impresso (papel e caneta). Portanto, o formulário *on-line* utilizado neste estudo mostra-se promissor, para ser utilizado em pesquisas futuras, sobre avaliação da aptidão física.

CONSIDERAÇÕES FINAIS

Os autores envolvidos no estudo não possuem qualquer potencial conflito de interesse.

CONTRIBUIÇÕES

D. A. Pereira contribuiu na concepção e planejamento do projeto, revisão intelectual crítica, coleta dos dados e aprovação da versão final, a ser publicada. J. L. Correia Júnior contribuiu na redação e elaboração do artigo, coleta dos dados, revisão intelectual crítica e aprovação da versão final, a ser publicada. H. D. Sobral contribuiu na redação e elaboração do artigo, coleta dos dados, revisão intelectual crítica e aprovação da versão final, a ser publicada. J. C. Albuquerque Costa contribuiu na redação e elaboração do artigo, na revisão intelectual crítica e aprovação da versão final, a ser publicada. M. P. Gonçalves contribuiu na redação e elaboração do artigo, revisão intelectual crítica e aprovação final, a ser publicada. R. F. Dias contribuiu na concepção e

planejamento do projeto, redação e elaboração do artigo, revisão intelectual crítica, análise, interpretação de dados, aprovação final, a ser publicada; e assumiu a responsabilidade por todos os aspectos do trabalho, na garantia da exatidão e integridade de qualquer parte da obra.

REFERENCIAS

1. Álvarez-Gallardo, I. C., Soriano-Maldonado, A., Segura-Jiménez, V., Carbonell-Baeza, A., Estévez-López, F., McVeigh, J. G., Delgado-Fernández, M., & Ortega, F. B. (2016). International Fitness Scale (IFIS): Construct Validity and Reliability in Women With Fibromyalgia: The al-Ándalus Project. *Archives of Physical Medicine and Rehabilitation*, 97(3), 395–404. <https://doi.org/10.1016/j.apmr.2015.08.416>
2. American College of Sports Medicine. (2017). *ACSM's Health-Related Physical Fitness Assessment* (5th ed.).
3. Belisario, J. S. M., Huckvale, K., Saje, A., Porcnik, A., Morrison, C. P., & Car, J. (2015). Comparison of self administered survey questionnaire responses collected using mobile apps versus other methods. *Cochrane Database of Systematic Reviews*, 7. <https://doi.org/10.1002/14651858.MR000042.pub2>
4. Bicen, A., Uzunboylu, H., & Burgul, N. S. (2020). Evaluation of Participants' Opinions on Online Physical Fitness Training. *Revista de Psicología Del Deporte (Journal of Sport Psychology)*, 29(4), 25–32.
5. Braekman, E., Berete, F., Charafeddine, R. agreement of the self-administered questionnaire of the B. H. I. S. P. versus web-based mode, Demarest, S., Drieskens, S., Gisle, L., Molenberghs, G., Tafforeau, J., Van der Heyden, J., & Van Hal, G. (2018). Measurement agreement of the self-administered questionnaire of the Belgian Health Interview Survey: Paper-and-pencil versus web-based mode. *PLoS One*, 13(5), e0197434. <https://doi.org/10.1371/journal.pone.0197434>
6. Brown, T. A. (2015). *Confirmatory factor*

Parâmetros psicométricos do IFIS – BRA

- analysis for applied research* (2th ed.). The Guilford Press.
7. Buchanan, T., Ali, T., Heffernan, T. M., Ling, J., Parrott, A. C., Rodgers, J., & Scholey, A. B. (2005). Nonequivalence of on-line and paper-and-pencil psychological tests: The case of the prospective memory questionnaire. *Behavior Research Methods*, 37(1), 148–154. <https://doi.org/10.3758/BF03206409>
 8. Campos, J. A. D. B., Zucoloto, M. L., Bonafé, F. S. S., Jordani, P. C., & Maroco, J. (2011). Reliability and validity of self-reported burnout in college students: A cross randomized comparison of paper-and-pencil vs. online administration. *Computers in Human Behavior*, 27(5), 1875–1883. <https://doi.org/10.1016/j.chb.2011.04.011>
 9. Chirivella, E. C. (2016). La especialidad profesional en Psicología del Deporte. *Revista de Psicología Aplicada Al Deporte y El Ejercicio Físico*, 1(1), 1–12. <https://doi.org/10.5093/rpadef2016a2>
 10. Cohen, J. (1977). Statistical Power Analysis for the Behavioral Sciences. In *Lawrence Erlbaum Associates*. Elsevier. <https://doi.org/10.1016/C2013-0-10517-X>
 11. Damásio, B. F. (2013). Contribuições da Análise Fatorial Confirmatória Multigrupo (AFCMG) na avaliação de invariância de instrumentos psicométricos. *Psico-USF*, 18(2), 211–220. <https://doi.org/10.1590/S1413-82712013000200005>
 12. Donovan, M. A., Drasgow, F., & Probst, T. M. (2000). Does computerizing paper-and-pencil job attitude scales make a difference? New IRT analyses offer insight. *Journal of Applied Psychology*, 85(2), 305–313. <https://doi.org/10.1037/0021-9010.85.2.305>
 13. Ebert, J. F., Huibers, L., Christensen, B., & Christensen, M. B. (2018). Paper- or Web-Based Questionnaire Invitations as a Method for Data Collection: Cross-Sectional Comparative Study of Differences in Response Rate, Completeness of Data, and Financial Cost. *Journal of Medical Internet Research*, 20(1), e24. <https://doi.org/10.2196/jmir.8353>
 14. Español-Moya, M. N., & Ramírez-Vélez, R. (2014). Validación del cuestionario International Fitness Scale (IFIS) en sujetos colombianos de entre 18 y 30 años de edad. *Revista Española de Salud Pública*, 88(2), 271–278. <https://doi.org/10.4321/S1135-57272014000200009>
 15. Falconi, C. A., Zanetti, M. C., Santos, T. de A., Dias, H. M., Brandao, M. R. F., & Neves, A. N. (2019). Relação entre antropometria, gordura corporal e autoconceito de adolescentes do sexo feminino. *Cuadernos de Psicología Del Deporte*, 19(2), 256–264. <https://doi.org/10.6018/cpd.348001>
 16. Finegan, J. E., & Allen, N. J. (1994). Computerized and written questionnaires: Are they equivalent? *Computers in Human Behavior*, 10(4), 483–496. [https://doi.org/10.1016/0747-5632\(94\)90042-6](https://doi.org/10.1016/0747-5632(94)90042-6)
 17. Gelder, M. M. H. J. V., Bretveld, R. W., & Roeleveld, N. (2010). Web-based Questionnaires: The Future in Epidemiology? *American Journal of Epidemiology*, 172(11), 1292–1298. <https://doi.org/10.1093/aje/kwq291>
 18. George, D., & Mallery, P. (2003). *SPSS for Windows Step by Step: A Simple Guide and Reference* (4th ed.).
 19. González-Ruiz, S. L., Gómez-Gallego, I., Pastrana-Brincones, J. L., & Hernández-Mendo, A. (2015). Algoritmos de clasificación y redes neuronales en la observación automatizada de registros. *Cuadernos de Psicología Del Deporte*, 15(1), 31–40.
 20. González-Ruiz, S. L., Hernández-Mendo, A., & Pastrana Brincones, J. L. (2010). Herramienta software para la evaluación psicosocial de deportistas y entornos deportivos. *Lecturas: EF y Deportes. Revista Digital*, 15(144), 1–15.
 21. Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Análise multivariada de dados*. Bookman editora.
 22. Hedman, E., Ljótsson, B., Rück, C., Furmark, T., Carlbring, P., Lindefors, N., & Andersson, G. (2010). Internet administration of self-report measures commonly used in research on social anxiety disorder: A

- psychometric evaluation. *Computers in Human Behavior*, 26(4), 736–740. <https://doi.org/10.1016/j.chb.2010.01.010>
23. Hirao, N., Koizumi, K., Ikeda, H., & Ohira, H. (2021). Reliability of Online Surveys in Investigating Perceptions and Impressions of Faces. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.733405>
 24. Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., Garnier-Villarreal, M., Selig, J., Boulton, A., & Preacher, K. (2021). *Package 'semTools.'*
 25. Joubert, T., & Kriek, H. J. (2009). Psychometric comparison of paper-and-pencil and online personality assessments in a selection setting. *SA Journal of Industrial Psychology*, 35(1), 1. <https://doi.org/10.4102/sajip.v35i1.727>
 26. King, W. C., & Miles, E. W. (1995). A quasi-experimental assessment of the effect of computerizing noncognitive paper-and-pencil measurements: A test of measurement equivalence. *Journal of Applied Psychology*, 80(6), 643–651. <https://doi.org/10.1037/0021-9010.80.6.643>
 27. Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
 28. Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
 29. Lopes, V. P., & Sá, C. (2021). The use of physical activity trackers devices and physical activity levels in adolescents and adults. *Cuadernos de Psicología Del Deporte*, 21(1), 258–270.
 30. Marôco, J. (2010). *Análise De Equações Estruturais - Fundamentos Teóricos, Software & Aplicações.*
 31. Marôco, J. (2014). *Análise de Equações Estruturais: Fundamentos Teóricos, Software & Aplicações.* In *Pêro Pinheiro: ReportNumber.*
 32. Mead, A. D., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin*, 114(3), 449–458. <https://doi.org/10.1037/0033-2909.114.3.449>
 33. Muthén, L. K., & Muthén, B. O. (2014). *Mplus User's Guide* (7th ed.).
 34. Ortega, A. M. S., & Montero, F. J. O. (2021). Relación entre resiliencia y rendimiento en deportistas. Revisión sistemática. *Revista de Psicología Aplicada Al Deporte y Al Ejercicio Físico*, 6(2), 1–11.
 35. Ortega, Ruiz, J. R., España-Romero, V., Vicente-Rodriguez, G., Martínez-Gómez, D., Manios, Y., Béghin, L., Molnar, D., Widhalm, K., Moreno, L. A., Sjöström, M., & Castillo, M. J. (2011). The International Fitness Scale (IFIS): usefulness of self-reported fitness in youth. *International Journal of Epidemiology*, 40(3), 701–711. <https://doi.org/10.1093/ije/dyr039>
 36. Pans, M., Úbeda-Colomer, J., & Devís-Devís, J. (2021). Validación de la Athletic Identity Measurement Scale en Estudiantes Universitarios con Discapacidad y Diferencias según Variables Sociodemográficas. *Revista de Psicología Del Deporte (Journal of Sport Psychology)*, 30(2), 1–10.
 37. Pereira, D. A., Carvas Junior, N., Correia Júnior, J. L., & Freitas-Dias, R. (2019). Transcultural adaptation: translation and validation of content of the brazilian version of the the international fitness scale instrument. *Adolescência & Saúde*, 16(3), 80–92.
 38. Pereira, D. A., Carvas Junior, N., Correia Júnior, J. L., & Freitas-Dias, R. (2021). *IFIS - versão brasileira* (p. 1). <http://profith.ugr.es/pages/investigacion/recur-sos/ifis/ifisversaiobrasileira>
 39. Pereira, D. A., Correia Júnior, J. L., Carvas Junior, N., & Freitas-Dias, R. de. (2020). Reliability of questionnaire The International Fitness Scale: a systematic review and meta-analysis. *Einstein (São Paulo)*, 18, 1–9. https://doi.org/10.31744/einstein_journal/2020r5232
 40. Polit, D. F. (2014). Getting serious about test–retest reliability: a critique of retest research and some recommendations. *Quality*

Parâmetros psicométricos do IFIS – BRA

- of Life Research*, 23(6), 1713–1720. <https://doi.org/10.1007/s11136-014-0632-9>
41. Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
42. Ramírez-Vélez, R., Cruz-Salazar, S. M., Martínez, M., Cadore, E. L., Alonso-Martínez, A. M., Correa-Bautista, J. E., Izquierdo, M., Ortega, F. B., & García-Hermoso, A. (2017). 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. *PeerJ*, 5. <https://doi.org/10.7717/peerj.3351>
43. Reigal, R. E., Pastrana-Brincones, J. L., González-Ruiz, S. L., Hernández-Mendo, A., Morillo-Baro, J. P., & Morales-Sánchez, V. (2020). Use of Data Mining to Determine Usage Patterns of an Online Evaluation Platform During the COVID-19 Pandemic. *Frontiers in Psychology*, 11, 2579. <https://doi.org/10.3389/fpsyg.2020.588843>
44. Rosseel, Y. (2014). The lavaan tutorial. *Department of Data Analysis: Ghent University*.
45. Souza, A. C., Alexandre, N. M. C., & Guirardello, E. de B. (2017). Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. *Epidemiologia e Serviços de Saúde*, 26(3), 649–659. <https://doi.org/10.5123/S1679-49742017000300022>
46. Stawarz, K., Preist, C., Tallon, D., Wiles, N., & Coyle, D. (2018). User Experience of Cognitive Behavioral Therapy Apps for Depression: An Analysis of App Functionality and User Reviews. *Journal of Medical Internet Research*, 20(6), e10120. <https://doi.org/10.2196/10120>
47. Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics, 6th (ed.)* Boston, MA. Pearson/Allyn & Bacon.
48. World Health Organization. (2020). *WHO guidelines on physical activity and sedentary behaviour: at a glance*.
49. Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research, and Evaluation*, 12(1), 3.
50. Zeleke, A. A., Worku, A. G., Demissie, A., Otto-Sobotka, F., Wilken, M., Lipprandt, M., Tilahun, B., & Röhrig, R. (2019). Evaluation of Electronic and Paper-Pen Data Capturing Tools for Data Quality in a Public Health Survey in a Health and Demographic Surveillance Site, Ethiopia: Randomized Controlled Crossover Health Care Information Technology Evaluation. *JMIR MHealth and UHealth*, 7(2), e10995. <https://doi.org/10.2196/10995>
51. Zhang, F., Huang, L., & Peng, L. (2020). The Degree of Influence of Daily Physical Activity on Quality of Life in Type 2 Diabetics. *Frontiers in Psychology*, 11, 1292. <https://doi.org/10.3389/fpsyg.2020.01292>