



## TEORIA DA RESPOSTA AO ITEM NAS PESQUISAS EM SAÚDE PÚBLICA

### THEORY OF THE RESPONSE TO THE ITEM IN RESEARCH IN PUBLIC HEALTH

### TEORÍA DE LA RESPUESTA AL ÍTEM EN LAS INVESTIGACIONES EN SALUD PÚBLICA

Diego Eller Gomes<sup>1</sup>, José Luís Guedes dos Santos<sup>2</sup>, José Wicto Pereira Borges<sup>3</sup>, Murilo Pedroso Alves<sup>4</sup>, Dalton Francisco de Andrade<sup>5</sup>, Alacoque Lorenzini Erdmann<sup>6</sup>

#### RESUMO

**Objetivo:** analisar a utilização da Teoria da Resposta ao Item (TRI) nas pesquisas em saúde pública. **Método:** revisão integrativa, sem delimitação temporal, nas bases de dados LILACS, SCOPUS, MEDLINE/PUBMED e na biblioteca virtual SCIELO. Adotaram-se as etapas de leitura, contagem de ocorrência de características das publicações e organização em categorias temáticas no processo de análise. **Resultados:** 56 estudos atenderam aos critérios de inclusão que buscaram avaliar a validade e a confiabilidade de instrumentos de mensuração e realizar ajustes psicométricos. Destacaram-se, como traços latentes mais avaliados, a qualidade de vida, a saúde de idosos e o letramento em saúde. **Conclusão:** evidenciam-se as possibilidades de utilização da TRI para a avaliação de uma diversidade de traços antes relacionados à saúde pública ao permitir a transformação de resultados subjetivos em probabilidades. O desenvolvimento de novos instrumentos pode considerar a avaliação de traços latentes relevantes para a saúde pública, como a integralidade da assistência na atenção básica, o vínculo do usuário aos serviços de saúde e perspectivas sobre vulnerabilidades, contribuindo para o avanço do conhecimento científico. **Descritores:** Métodos e Procedimentos Estatísticos; Inquéritos e Questionários; Análise Quantitativa; Saúde Pública; Enfermagem; Avaliação em Enfermagem.

#### ABSTRACT

**Objective:** to analyze the use of Item Response Theory (IRT) in public health research. **Method:** integrative review, without temporal delimitation, in the databases LILACS, SCOPUS, MEDLINE / PUBMED and in the SCIELO virtual library. We adopted the steps of reading, counting of occurrence of characteristics of the publications and organization in thematic categories in the process of analysis. Results: 56 studies met the inclusion criteria that sought to evaluate the validity and reliability of measurement instruments and to make psychometric adjustments. The quality of life, the health of the elderly and the health literacy were highlighted as more evaluated latent traits. Conclusion: the possibilities of using IRT for the evaluation of a diversity of traits previously related to public health are evidenced by allowing the transformation of subjective results into probabilities. The development of new instruments may consider the evaluation of latent traits relevant to public health, such as the integrality of care in basic care, the user's link to health services and perspectives on vulnerabilities, contributing to the advancement of scientific knowledge. **Descriptors:** Statistical Methods and Procedures; Surveys and Questionnaires; Quantitative Analysis; Public Health; Nursing; Nursing Assessment.

#### RESUMEN

**Objetivo:** analizar la utilización de la Teoría de la Respuesta al Ítem (TRI) en las investigaciones en salud pública. **Método:** revisión integrativa, sin delimitación temporal, en las bases de datos LILACS, SCOPUS, MEDLINE / PUBMED y en la biblioteca virtual SCIELO. Se adoptaron las etapas de lectura, conteo de ocurrencia de características de las publicaciones y organización en categorías temáticas en el proceso de análisis. **Resultados:** 56 estudios atendieron a los criterios de inclusión, que buscaron evaluar la validez y la confiabilidad de instrumentos de medición y realizar ajustes psicométricos. Se destacaron como rasgos latentes más evaluados, la calidad de vida, la salud de ancianos y el letramiento en salud. **Conclusión:** se evidencian las posibilidades de utilización de la TRI para la evaluación de una diversidad de rasgos latentes relacionados a la salud pública, al permitir la transformación de resultados subjetivos en probabilidades. El desarrollo de nuevos instrumentos puede considerar la evaluación de rasgos latentes relevantes para la salud pública, como la integralidad de la asistencia en la atención básica, el vínculo del usuario a los servicios de salud y perspectivas sobre vulnerabilidades, contribuyendo al avance del conocimiento científico. **Descritores:** Métodos y Procedimientos Estadísticos; Encuestas y Questionarios; Análisis Cuantitativo; Salud Pública; Enfermería; Evaluación em Enfermería.

<sup>1</sup>Doutorando em Engenharia de Produção, Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina/UFSC, Florianópolis (SC), Brasil. E-mail: <https://orcid.org/0000-0003-4434-5492>; <sup>2</sup>Doutor, Programa de Pós-Graduação em Enfermagem, Universidade Federal de Santa Catarina/UFSC, Florianópolis (SC), Brasil. E-mail: [joseenfermagem@gmail.com](mailto:joseenfermagem@gmail.com); ORCID iD: <https://orcid.org/0000-0003-3186-8286>; <sup>3</sup>Doutor, Programa de Pós-Graduação em Saúde e Comunidade, Universidade Federal do Piauí/UFPI, Floriano (PI), Brasil. E-mail: [josewictoborges@gmail.com](mailto:josewictoborges@gmail.com); ORCID iD: <https://orcid.org/0000-0002-3292-1942>; <sup>4</sup>Doutorando, Programa de Pós-Graduação em Enfermagem, Universidade Federal de Santa Catarina/UFSC, Florianópolis (SC), Brasil. E-mail: [murilopedrosoalves@gmail.com](mailto:murilopedrosoalves@gmail.com); ORCID iD: <https://orcid.org/0000-0002-2704-6848>; <sup>5</sup>Doutor, Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina/UFSC, Florianópolis (SC), Brasil. E-mail: [dalton.andrade@ufsc.br](mailto:dalton.andrade@ufsc.br); ORCID iD: <https://orcid.org/0000-0002-4403-980X>; <sup>6</sup>Doutora, Programa de Pós-Graduação em Enfermagem, Universidade Federal de Santa Catarina/UFSC, Florianópolis (SC), Brasil. E-mail: [alacoque@newsite.com.br](mailto:alacoque@newsite.com.br); ORCID iD: <https://orcid.org/0000-0003-3964-1115>

## INTRODUÇÃO

Considera-se o uso de instrumentos e testes psicométricos como um importante avanço na área de avaliação em saúde por possibilitar a investigação de diversos fenômenos.<sup>1</sup> Deve-se a isso o fato de que os testes permitem o estabelecimento de um referencial e, conseqüentemente, reduzem os vieses subjetivos das avaliações.<sup>2</sup>

Iniciou-se, no século XIX, o desenvolvimento de instrumentos de avaliação psicológica, paralelo ao avanço da ciência positivista e à identificação da importância de medidas objetivas e válidas para a pesquisa clínica.<sup>1</sup> Surgiu, a partir disso, a necessidade de teorias que avaliassem as propriedades psicométricas dos instrumentos de mensuração.<sup>1,3</sup>

Nesse contexto, identificam-se duas grandes linhas teóricas da psicometria: a Teoria Clássica dos Testes (TCT) e a Teoria da Resposta ao Item (TRI). Tem-se, como foco da TCT, o próprio instrumento de medida com o objetivo de determinar as propriedades ou os parâmetros métricos do teste.<sup>4</sup> Já o foco da TRI concentra-se no estudo individualizado dos itens componentes de um grupo - teste ou banco de itens - ao contrário da sua antecessora.<sup>3</sup>

A TRI, expressão da moderna psicometria, constitui-se em um grupo de modelos matemáticos que permitem avanços nos estudos de validade e confiabilidade de instrumentos a partir de análises robustas aplicadas aos diversos tipos de instrumentos de mensuração. Dentre os modelos da TRI, incluem-se: modelo de resposta nominal; modelo normal modificado; modelo de eleição múltipla; modelo de escala de estimação; modelo de crédito parcial; modelo de crédito parcial generalizado; modelo de resposta gradual e modelo contínuo.<sup>5</sup>

Teve-se em vários países, assim como no Brasil, a TRI amplamente divulgada e utilizada na avaliação de testes educacionais,<sup>6-7</sup> de gestão da qualidade,<sup>8</sup> *marketing*,<sup>9</sup> entre outros.<sup>10</sup> Utilizam-se, na área da saúde, alguns dos modelos da TRI para a construção,<sup>11</sup> validação<sup>12-3</sup> e avaliação de instrumentos e programas.<sup>14</sup> Ressalta-se a importância da saúde pública como ciência que fomenta o desenvolvimento de estratégias e ferramentas para melhorar a qualidade de vida da população por meio da gestão do cuidado e dos serviços de saúde, da prevenção de doenças e agravos à saúde da doença, da promoção da saúde e da educação do paciente, família e coletividade. Contribui-se

com a prática dos diversos profissionais da saúde, especialmente da Enfermagem.

Ao se considerar as possibilidades de utilização da TRI, o desenvolvimento de um estudo visando à identificação de quais são e como vêm sendo utilizados os modelos matemáticos da TRI na área da saúde pública apresenta-se como uma contribuição para pesquisadores interessados em utilizá-la em suas investigações. Até este momento, não há uma síntese do conhecimento que possibilite analisar o panorama e as perspectivas do uso da TRI na área da saúde pública. Apresenta-se, portanto, um avanço no conhecimento científico relacionado à temática. Além disso, uma análise nesse sentido poderá mostrar perspectivas para avanços aos estudos de mensuração do comportamento em saúde. Assim, questiona-se: como tem sido utilizada a TRI nas pesquisas em saúde pública?

## OBJETIVO

- Analisar a utilização da Teoria da Resposta ao Item (TRI) nas pesquisas em saúde pública.

## MÉTODO

Revisão integrativa a partir da utilização de um instrumento de intervenção denominado *Knowledge Development Process - Constructivist (Proknow-C)*.<sup>15</sup> Esse instrumento subsidiou o processo de formação do portfólio bibliográfico (PB) referente ao fragmento da literatura: utilização da TRI nas pesquisas em saúde pública.

Tem-se a primeira etapa do *Proknow-C* com o objetivo de identificar as publicações científicas relativas ao tema definido pelos pesquisadores. Essa etapa constitui-se por três subetapas: (1) seleção do banco de artigos bruto; (2) filtragem do banco de artigos e (3) realização do teste de representatividade dos artigos. Em cada subetapa, os pesquisadores necessitam fazer escolhas resultando na seleção do portfólio, que é um conjunto restrito e relevante de artigos científicos representativos do fragmento da literatura sob investigação.

Adotaram-se como critérios para a inclusão dos estudos na revisão: explicitar a utilização da TRI em pesquisas na área da saúde pública a partir da combinação (“*item response theory*” AND “*public health*”) no título, no resumo ou nas palavras-chave; estar nos idiomas português, inglês ou espanhol; estar disponível *online*. Optou-se por não estabelecer delimitação temporal tendo em vista a pequena quantidade de estudos publicados até o ano de 2010. Aplicaram-se

Gomes DE, Santos JLG dos, Borges JWP et al.

Teoria da resposta ao item nas pesquisas...

esses critérios às bases de dados LILACS e SCOPUS, à biblioteca virtual *Scientific Electronic Library Online* (SCIELO) e ao Sistema Online de Busca e Análise de Literatura Médica (*Medical Literature Analysis and Retrieval System Online*) via PUBMED.

Apresenta-se, na figura 1, o processo de aplicação do *Proknow-C* para a seleção do portfólio de artigos representativos do fragmento da literatura.

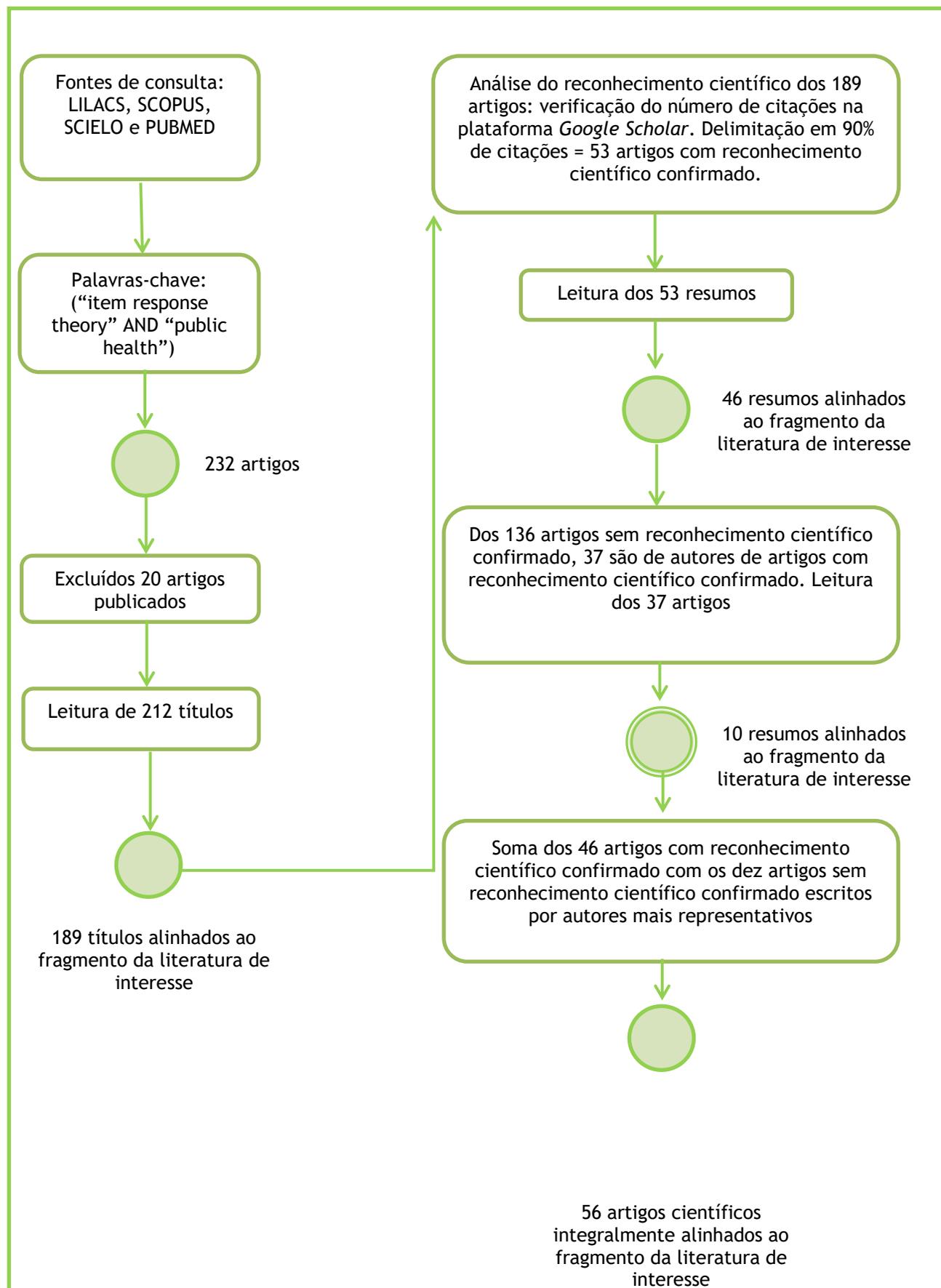


Figura 1. Etapas do processo de formação do portfólio bibliográfico. Florianópolis (SC), Brasil, 2018.

Constituiu-se o portfólio bibliográfico com 56 artigos científicos disponíveis *online* e alinhados integralmente ao tema da pesquisa.

Após a formação do portfólio de artigos científicos, procedeu-se à contagem de algumas características das publicações do fragmento da literatura investigado a fim de construir conhecimento nos pesquisadores e, posteriormente, na comunidade científica.<sup>15-6</sup> Definiram-se como características a serem analisadas: (1) quem são os pesquisadores com trajetória na área de conhecimento; (2) quais são os periódicos que têm dedicado espaço

para a publicação do assunto e (3) quais são os artigos mais citados pela comunidade científica, por meio da contagem de citações realizada na plataforma *Google Scholar*, em 25 de agosto de 2017.

Na última etapa, referente à análise de conteúdo, realizou-se uma análise crítica do PB selecionado descrevendo-se os dados e agrupando-os por similaridade. Para tanto, estabeleceram-se categorias *a priori*, tais como: objetivos dos estudos, traços latentes avaliados e modelos da TRI utilizados. Ao final das três etapas, identificou-se onde e como é

possível intervir cientificamente em relação ao tema investigado e justificar essas escolhas.

## RESULTADOS

Destacaram-se dois autores pelo maior número de publicações científicas utilizando a TRI em pesquisas relacionadas à saúde pública, ambos com 11 publicações. Um desses autores é professor das áreas de Direito, Política e Gestão em Saúde na Universidade de Medicina de Xangai, na China. O outro autor de destaque é professor de Políticas e Gestão em Saúde na Escola de Saúde Pública da Universidade de Boston, nos Estados Unidos, cujas pesquisas são voltadas à deficiência, fisioterapia e gerontologia.

O *Quality of Life Research* (sete artigos), *Archives of Physical Medicine and Rehabilitation* (seis publicações) e *Journal of Clinical Epidemiology* (três publicações) destacaram-se como os três periódicos com maior espaço para publicações usando a TRI em pesquisas na área da saúde pública. O periódico *Quality of Life Research* é

multidisciplinar, dedicado à comunicação de pesquisas originais, artigos teóricos e relatos metodológicos relacionados ao campo da qualidade de vida, em todas as ciências da saúde, com fator de impacto de 2,344. O periódico *Archives of Physical Medicine and Rehabilitation* possui enfoque na criação e uso do conhecimento no processo de reabilitação, por meio de pesquisas originais e relatórios clínicos sobre tendências na reabilitação médica, cujo fator de impacto é de 3.289. O terceiro periódico que mais tem dedicado espaço às publicações que utilizam a TRI na área da saúde pública - *Journal of Clinical Epidemiology* - visa a promover a qualidade da pesquisa de cuidados de saúde clínica por meio da promoção e aplicação de métodos inovadores, com especial atenção à formação de novos cientistas e líderes de prática clínica, e seu fator de impacto é 4.978.

Quanto aos artigos em destaque, a figura 2 apresenta os dez artigos mais citados pela comunidade científica, com ênfase para o estudo *Development of an easy-to-use Spanish Health Literacy Test*,<sup>17</sup> com 170 citações.

Título	Periódico	Ano	Primeiro autor	Nº de citações
Development of an easy-to-use Spanish Health Literacy test	Health Services Research	2006	Lee <sup>17</sup>	170
The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ)	BMC Public Health	2013	Osborne <sup>18</sup>	156
A two-tier full-information item factor analysis model with applications	Psychometrika	2010	Cai <sup>19</sup>	131
Development of a Japanese version of the Epworth Sleepiness Scale (JESS) based on item response theory	Sleep Medicine	2009	Takegami <sup>20</sup>	110
The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances	Quality of Life Research	2014	Ravens-Sieberer <sup>21</sup>	100
Short Assessment of Health Literacy-Spanish and English: a comparable test of health literacy for Spanish and English speakers	Health Services Research	2010	Lee <sup>22</sup>	73
Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers	BMC Medical Research Methodology	2012	Stochl <sup>23</sup>	65
Unidimensionality and bandwidth in the Center for Epidemiologic Studies Depression (CES-D) Scale	Journal of Personality Assessment	2006	Stansbury <sup>24</sup>	63
The development of the S-QoL 18: a shortened quality of life questionnaire for patients with schizophrenia	Schizophrenia Research	2010	Boyer <sup>25</sup>	57
Can item response theory reduce patient burden when measuring health status in neurological disorders? Results from Rasch analysis of the SF-36 physical functioning scale (PF-10)	Journal of Neurology, and Psychiatry	2001	Jenkinson <sup>26</sup>	48

Figura 2. Artigos mais citados pela comunidade científica que utilizaram a TRI na área da saúde pública. Florianópolis (SC), Brasil, 2018.

Nota-se que um artigo publicado em 2006 obteve o maior número de citações. No entanto, artigos publicados em 2013 e 2014 também figuram entre os mais citados, com

156 e 100 citações, respectivamente. Os artigos apresentados na tabela 1 são importantes fontes de pesquisa por evidenciar maior reconhecimento científico acerca da

aplicação da TRI em pesquisas na área da saúde pública.

Apresentam-se objetivos diversificados, no entanto, pode-se observar maior ocorrência

dos verbos avaliar, desenvolver e validar. Na figura 3, evidenciam-se as características dos estudos quanto aos objetivos, traços latentes e modelos da TRI utilizados.

Primeiro autor	Ano	Objetivo	Traço latente	Modelos da TRI
Jenkinson <sup>26</sup>	2001	Validar um construto	Função física em doença motora neurológica e Parkinson	ML1P <sup>b</sup>
Gulliford <sup>27</sup>	2004	Avaliar confiabilidade e validade	Segurança alimentar familiar	ML2P <sup>c</sup>
Lee <sup>17</sup>	2006	Desenvolver uma medida	Letramento em saúde para adultos	ML1P, ML2P, ML3P <sup>d</sup>
Metz <sup>28</sup>	2006	Avaliar um ajuste psicométrico	Qualidade de vida de pacientes com asma	MCPG <sup>e</sup>
Stansbury <sup>24</sup>	2006	Avaliar um ajuste psicométrico	Depressão	ML1P
Filiatrault <sup>29</sup>	2007	Avaliar confiabilidade e validade	Equilíbrio para atividades em idosos	MRG <sup>f</sup>
Jette <sup>30</sup>	2008	Desenvolver uma medida	Uso de pesquisas sobre envelhecimento	DIF <sup>g</sup>
Tao <sup>31</sup>	2008	Desenvolver uma medida	Mobilidade, atividades diárias e habilidades cognitivas	Não especifica
Allen <sup>32</sup>	2009	Validar um construto	Resultados pediátricos	ML1P
Forjaz <sup>33</sup>	2009	Avaliar confiabilidade e validade	Sintomas autônomos na doença de Parkinson	ML1P; DIF
Forjaz <sup>34</sup>	2010	Avaliar um ajuste psicométrico	Ansiedade e depressão hospitalar	ML1P; DIF
Haley <sup>35</sup>	2009	Avaliar um banco de itens	Função física em crianças com paralisia cerebral	ML1P; DIF
Haley <sup>36</sup>	2009	Avaliar um ajuste psicométrico	Função física em crianças com paralisia cerebral	ML2P; DIF
Jette <sup>37</sup>	2009	Desenvolver uma medida	Função física em pessoas com osteoartrite	ML2P
Takegami <sup>20</sup>	2009	Avaliar confiabilidade e validade	Sonolência	MCPG
Boyer <sup>25</sup>	2010	Avaliar um ajuste psicométrico	Qualidade de vida	MCPG; DIF
Brodersen <sup>38</sup>	2010	Avaliar um ajuste psicométrico	Consequências da triagem no câncer de pulmão	MCPG
Cai <sup>19</sup>	2010	Validar um construto	Qualidade de vida	ML2P
Cherepanov <sup>39</sup>	2010	Avaliar um ajuste psicométrico	Qualidade de vida	ML2P
Conrad <sup>40</sup>	2010	Avaliar confiabilidade e validade	Crime e violência	ML1P
Perron <sup>41</sup>	2010	Validar um construto	Transtornos pelo uso/abuso de inalantes	ML2P
Villaveces <sup>42</sup>	2010	Desenvolver uma medida	Prevenção de lesões e violência	ML1P, ML2P, ML3P
Conrad <sup>43</sup>	2011	Avaliar confiabilidade e validade	Abuso psicológico em idosos	ML1P; DIF
Dregan <sup>44</sup>	2011	Desenvolver uma medida	Códigos médicos	ML2P
Fang <sup>45</sup>	2011	Avaliar um ajuste psicométrico	Qualidade de vida de pessoas com deficiência	DIF
Riley <sup>46</sup>	2011	Avaliar um ajuste psicométrico	Saúde mental	MRG; DIF
Bann <sup>47</sup>	2012	Desenvolver uma medida	Vigilância de saúde pública de bem-estar	DIF
Chan <sup>48</sup>	2012	Avaliar equivalência de itens	Envelhecimento	ML2P
Fang <sup>49</sup>	2012	Avaliar um ajuste psicométrico	Qualidade de vida de idosos	MCPG; DIF
Jette <sup>50</sup>	2012	Calibrar um banco de itens	Ferimento funcional da medula espinhal	MRG; DIF
Stochl <sup>23</sup>	2012	Desenvolver um instrumento	Saúde e bem-estar	Mokken
Lee <sup>51</sup>	2013	Desenvolver um instrumento	Letramento em saúde oral	DIF
Lee <sup>22</sup>	2010	Desenvolver uma medida	Letramento em saúde oral para adultos	ML1P, ML2P, ML3P
Marfeo <sup>52</sup>	2013	Avaliar um ajuste psicométrico	Saúde comportamental	MRG; DIF
Marfeo <sup>53</sup>	2013	Validar um construto	Comportamento de saúde no trabalho	ML1P, ML2P, ML3P
Mcdonough <sup>54</sup>	2013	Avaliar um ajuste psicométrico	Reabilitação ambulatorial	CAT <sup>h</sup>

Osborne <sup>18</sup>	2013	Desenvolver uma medida	Letramento em saúde	MRG
Parsons <sup>55</sup>	2013	Avaliar confiabilidade e validade	Desordem hipersexual	MRG
Waller <sup>56</sup>	2013	Desenvolver uma medida	Conhecimento sobre HPV	ML1P, ML2P, ML3P
Bevans <sup>57</sup>	2014	Calibrar um banco de itens	Qualidade de vida	CAT
Hahn <sup>58</sup>	2014	Desenvolver uma medida	Saúde social em doenças crônicas	ML2P; DIF
Onishi <sup>59</sup>	2014	Desenvolver uma medida	Sintomas de ascite na cirrose	MCPG
Ravens-Sieberer <sup>21</sup>	2014	Desenvolver um teste adaptativo	Qualidade de vida relacionada à saúde da criança	MCPG
Rose <sup>60</sup>	2014	Calibrar um banco de itens	Função física	MRG
Spiegel <sup>61</sup>	2014	Desenvolver uma medida	Sintomas gastrointestinais	MRG
Zhong <sup>62</sup>	2014	Avaliar confiabilidade e validade	Escala de depressão	DIF
Chan <sup>63</sup>	2015	Avaliar um ajuste psicométrico	Desempenho cognitivo no envelhecimento	MRG
Epstein <sup>64</sup>	2015	Avaliar confiabilidade e validade	Impacto de educação em saúde	MRG; ML2P; DIF
Gerrard <sup>65</sup>	2015	Avaliar confiabilidade e validade	Integração de adultos feridos por queimadura	Não especificado
Honda <sup>66</sup>	2015	Desenvolver uma medida	Qualidade de vida em cirurgia gastrointestinal	MCPG
Michel <sup>67</sup>	2015	Desenvolver uma medida	Qualidade de vida em pacientes com esclerose múltipla	DIF
Schalet <sup>68</sup>	2015	Desenvolver uma medida	Saúde física e mental	MRG
Teicher <sup>69</sup>	2015	Desenvolver uma medida	Exposição a maus-tratos e abuso	ML1P
Coster <sup>70</sup>	2016	Avaliar um ajuste psicométrico	Mobilidade, fadiga e dor em crianças com paralisia cerebral	MRG; DIF
Mcdonough <sup>71</sup>	2016	Avaliar um ajuste psicométrico	Função cognitiva na reabilitação ambulatorial	CAT
Marino <sup>72</sup>	2017	Desenvolver um conjunto de itens	Impacto social das queimaduras em adultos	CAT

Figura 3. Objetivos, traços latentes e modelos da TRI utilizados nos estudos na área da saúde pública. Florianópolis (SC), Brasil, 2018.

b: Modelo logístico de um parâmetro

c: Modelo logístico de dois parâmetros

d: Modelo logístico de três parâmetros

e: Modelo de crédito parcial generalizado

f: Modelo de resposta gradual

g: Funcionamento diferencial do item

h: Teste adaptativo computadorizado

## DISCUSSÃO

Dentre os periódicos mais representativos, constatou-se que o *Journal of Clinical Epidemiology* busca evidenciar a aplicação de métodos inovadores na área da saúde, enquanto o periódico *Quality of Life Research* vem dedicando espaço para publicações relacionadas à qualidade de vida sendo este o traço latente mais avaliado nas pesquisas da saúde pública que utilizam a TRI.<sup>19,28,32,35,41,45,56,59,66-7</sup> Portanto, a avaliação da qualidade de vida, por meio da utilização dos modelos matemáticos e estatísticos da TRI, vem ganhando, cada vez mais, espaço nas publicações científicas e pode ser considerada uma combinação inovadora.

Quanto aos objetivos dos artigos, constatou-se que a TRI tem sido utilizada com três principais finalidades: avaliar instrumentos,<sup>27-9,32-3,35,39,41-2,44-5,51-2,62,65,70</sup> validar construtos<sup>28,31,40,54,37</sup> e desenvolver instrumentos.<sup>51,62,75</sup> Esses achados confirmam as possibilidades e os benefícios do uso da TRI

indicados em estudos clássicos.<sup>4,5,73</sup> Identificaram-se, também, objetivos diversificados, tais como: desenvolver a versão japonesa de uma escala de sonolência;<sup>20</sup> mensurar sintomas gastrointestinais relatados por pacientes;<sup>61</sup> explorar a viabilidade de um sistema de adaptação funcional em mobilidade, atividades diárias e habilidades cognitivas<sup>31</sup> e descrever o processo de calibração do índice de lesão funcional da medula espinhal.<sup>49</sup> Quatro estudos científicos buscaram harmonizar, em conjunto, os resultados de outras pesquisas.<sup>44,63,68</sup>

Com a finalidade de desenvolver ou avaliar instrumentos, bem como avaliar traços latentes, o modelo unidimensional de um, dois ou três parâmetros foi o mais recorrente e utilizado em 17 estudos analisados. O modelo logístico de três parâmetros (ML3)<sup>74</sup> é o mais completo, pois, além dos parâmetros de dificuldade e discriminação dos itens, considera, também, a probabilidade de acerto ao acaso quando o respondente possui um

baixo nível de aptidão ou traço latente.<sup>73</sup> Porém, somente cinco estudos adoram o ML3, geralmente em comparação aos modelos de um e dois parâmetros,<sup>17,38,48,51,55</sup> pois o acerto ao acaso parece não fazer tanto sentido quando comparado à sua aplicação na área educacional, que busca representar a probabilidade de um aluno com baixa habilidade responder corretamente ao item, ao acaso, ou chutar.<sup>6</sup>

Por meio do ML3, obtém-se facilmente o modelo logístico unidimensional de dois parâmetros (ML2),<sup>75</sup> por exemplo, quando não existe a possibilidade de acerto ao acaso. Dessa forma, esse modelo destacou-se como o mais utilizado nas pesquisas em saúde pública,<sup>19,28,30,35,37-38,40,44,48,51,55,57,64</sup> pois, além da dificuldade do item, contempla o parâmetro de discriminação dos itens que se refere à capacidade do item para discriminar os respondentes com diferentes níveis de aptidão ou traço latente.<sup>73</sup>

Dentre os modelos unidimensionais da TRI para itens dicotômicos, aqueles corrigidos como certo ou errado, pertence ou não pertence, satisfeito ou insatisfeito, entre outros, o modelo logístico de um parâmetro (ML1) ou modelo de Rasch é o mais simples, pois considera que todos os itens possuem o mesmo poder de discriminação e contempla, em sua fórmula, somente o parâmetro de dificuldade dos itens. Após o modelo de dois parâmetros da TRI, identificou-se o ML1 como o mais utilizado.<sup>17,26-8,31-34,,36,38-9,48,51,55,69</sup>

Em relação os modelos para itens politômicos, o modelo de resposta gradual (MRG) assume que as categorias de resposta de um item podem ser ordenadas entre si.<sup>76</sup> Nesses casos, os itens de múltipla escolha são avaliados de forma graduada, elaborados com categorias intermediárias ordenadas, em que se busca a obtenção de mais informações das respostas dos indivíduos e não apenas identificar se deram respostas corretas ou incorretas. Identificou-se a adoção desse modelo em 11 estudos que aplicaram a TRI em pesquisas na área da saúde pública.<sup>29,42,46,50,55,60-1,63-4,68,70</sup>

Outro achado evidenciado neste estudo refere-se à utilização do modelo da TRI de crédito parcial generalizado (MCPG) para itens politômicos nas pesquisas em saúde pública.<sup>77</sup> O modelo está relacionado a itens com respostas politômicas graduais da TRI e considera as probabilidades de escolha para cada uma das categorias de resposta como adotado em oito estudos científicos relacionados à saúde pública.<sup>28,33,45,58-9,66</sup> No modelo, o respondente ganha mais crédito (teria um maior valor para seu traço latente)

à medida que sua resposta se aproxima da resposta mais completa.<sup>78</sup>

Dentre os estudos selecionados, 19 deles realizaram a análise do funcionamento diferencial do item (DIF) com o objetivo de verificar se a probabilidade de acerto ou concordância com o item variou entre grupos de respondentes por sexo, idade, nível de escolaridade, localização geográfica e/ou nível de gravidade da doença.<sup>28-9,32,34,39,41-3,45-6,48,50,57,62,64,67,70</sup> A utilização do DIF na saúde pública e, conseqüentemente, por profissionais da Enfermagem, pode possibilitar a investigação de variáveis que influenciam as avaliações em saúde permitindo evidenciar as vulnerabilidades ou determinantes sociais da saúde.

A análise do DIF visa à detecção de itens cuja probabilidade de acerto difere entre grupos distintos cujos respondentes possuem o mesmo nível de traço latente na variável medida. Um bom modelo de medida deve requerer que a pessoa mais capaz, ou com maior nível de um traço latente, tenha sempre mais possibilidade de acerto em um item do que uma pessoa menos capaz. Toda pessoa deve ter mais possibilidade de acerto em um item fácil (ou comum) do que em um difícil (ou raro) e, finalmente, essas condições devem ser produto da pessoa e da posição do item no traço ou habilidade e não baseadas na raça, no sexo ou em outras características dos respondentes.<sup>79</sup>

Nos estudos, identificaram-se benefícios decorrentes da utilização da TRI em pesquisas na área da saúde pública tais como: criação de ferramentas de avaliação que possam ser aplicadas a populações de diferentes idiomas, no caso de pesquisas transculturais;<sup>57</sup> verificação da pertinência das categorias de resposta dos itens;<sup>33</sup> identificação de viés entre os grupos de respondentes em razão de suas características, por meio da análise do DIF;<sup>33-4,44,70</sup> elaboração de testes que não necessitam de muitos itens, pois a análise de seus parâmetros permitem a identificação dos itens com maior poder de discriminação, o que pode reduzir o erro e o tamanho da amostra;<sup>39,40-1,56</sup> alteração ou substituição de itens de forma iterativa, possibilitadas pela metodologia dos testes adaptativos computadorizados (CAT), em que a apresentação de um item se baseia nas informações obtidas das respostas aos itens aplicados anteriormente.<sup>29,42,60</sup>

De modo mais específico, alguns estudos apresentaram, como benefícios da TRI, a possibilidade de validação de medidas como a de autorrelato de abuso psicológico em idosos,<sup>36</sup> a avaliação e comparação entre o

nível de letramento em saúde de falantes de espanhol e inglês, aspecto positivo, inclusive, por subsidiar o processo de conscientização dos profissionais de Enfermagem em relação às necessidades dos pacientes com baixo nível de letramento em saúde, o que pode ajudar a reduzir a complexidade linguística utilizada na comunicação entre profissionais e pacientes.<sup>49</sup>

Evidenciam-se alguns cuidados a serem tomados ao desenvolver instrumentos de avaliação de traços latentes relacionados à saúde pública, como a calibração dos parâmetros dos itens a partir da aplicação do instrumento a pessoas com diferentes níveis de traço latente e a elaboração de itens com diferentes níveis de dificuldade reiterando-se a indicação apresentada em estudos clássicos acerca da utilização da TRI.<sup>5,73</sup>

Esta revisão tem como limitação o uso de delineamento descritivo em relação aos estudos que aplicaram a TRI em pesquisas relacionadas à saúde pública, sem incorrer em uma análise aprofundada dos traços latentes mais explorados em saúde pública como qualidade de vida, saúde de idosos e letramento em saúde. Além disso, alguns artigos não evidenciaram o modelo da TRI utilizado, o que dificultou a análise desses estudos.

## CONCLUSÃO

Utilizando-se a TRI, possibilita-se a avaliação de diversos construtos relevantes à saúde pública por meios matemáticos transformando-se os resultados subjetivos das avaliações em probabilidades. Para essa passagem, utilizou-se uma variedade de modelos, principalmente os modelos de um e dois parâmetros, a fim de determinar as qualidades de validade e confiabilidade dos instrumentos que buscaram avaliar, principalmente, os traços latentes qualidade de vida, saúde de idosos e letramento em saúde.

Trabalharam-se, também, os traços latentes, a partir do funcionamento diferencial de itens, para observar se determinadas características podem influenciar as respostas dos indivíduos. Assim, ao responder ao questionamento: como tem sido utilizada a TRI nas pesquisas em saúde pública, apresentam-se novas possibilidades de análise da confiabilidade e validade de instrumentos de mensuração em saúde pública. Tais perspectivas poderão ser utilizadas no estabelecimento de novos instrumentos considerando-se traços latentes relevantes para a saúde pública, como a integralidade da assistência na atenção básica, perspectivas sobre vulnerabilidades,

comportamento de promoção da saúde, vínculo do usuário e equidade no acesso aos serviços de saúde, contribuindo para futuras pesquisas na área da saúde pública e Enfermagem.

Evidencia-se, como campo fecundo para futuras investigações, o panorama teórico-analítico apresentado, especialmente, na área da Enfermagem. Denota-se a importância dos profissionais de Enfermagem na busca da emancipação dos usuários no contexto das discussões sobre os determinantes sociais de saúde condicionantes para o processo de viver saudável relacionando-se às vulnerabilidades socioeconômicas, culturais e ambientais.

## REFERENCES

1. Sartes LMA, Souza-Formigoni MLO. Advances in psychometrics: from Classical Test Theory to Item Response Theory. *Psicol Reflex Crit.* 2013;26(2):241-50. Doi: <http://dx.doi.org/10.1590/S0102-79722013000200004>
2. DeMars C. Item response theory: understanding statistics measurement. New York: Oxford University Press; 2010.
3. Andriola WB. *Psicometria Moderna: características e tendências.* Est Aval Educ. 2009 May/Aug;20(43):319-40. Doi: <http://dx.doi.org/10.18222/eae204320092052>
4. Pasquali L, Primi R. Basic theory of Item Response Theory - IRT. *Aval Psicol [Internet].* 2003 Dec [cited 2018 Feb 12];2(2):99-110. Available from: [http://pepsic.bvsalud.org/pdf/avp/v2n2/v2n2\\_a02.pdf](http://pepsic.bvsalud.org/pdf/avp/v2n2/v2n2_a02.pdf)
5. Embretson S, Reise S. Item response theory. Hoboken: Taylor and Francis; 2013.
6. Klein R. Use of Item Response Theory at the Brazilian Assessment System for Basic Education. *Rev Meta.* 2009;1(2):125-40. Doi: <http://dx.doi.org/10.22347/2175-2753v1i2.38>
7. Chen CM, Lee HM, Chen YH. Personalized e-learning system using item response theory. *Comput Educ.* 2005 Apr;44(3):237-55. Doi: <https://doi.org/10.1016/j.compedu.2004.01.006>
8. Alexandre JWC, Andrade DF, Vasconcelos AP, Araújo AMS. A construct analysis proposal for measuring the total quality management critical factors through the item response theory. *Gest Prod.* 2002 Aug;9(2):129-41. Doi: <http://dx.doi.org/10.1590/S0104-530X2002000200003>
9. Jong MG, Steenkamp JBEM, Fox JP, Baumgartner H. Using item response theory to measure extreme response style in marketing research: a global investigation. *J Mark Res.*

- 2008 Feb;45(1):104-15. Doi: <https://doi.org/10.1509/jmkr.45.1.104>
10. Pasquali L. *Psicometria*. Petrópolis: Vozes; 2009.
11. Rodrigues MTP, Moreira TMM, Andrade DF. Elaboration and validation of instrument to assess adherence to hypertension treatment. *Rev Saúde Pública*. 2014 Apr;48(2):232-9. Doi: <http://dx.doi.org/10.1590/S0034-8910.2014048005044>
12. César CC, Mambrini JVM, Ferreira FR, Lima-Costa MF. Functional capacity in the elderly: analyzing questions on mobility and basic and instrumental activities of daily living using Item Response Theory. *Cad Saúde Pública*. 2015 May;31(5):931-45. Doi: <http://dx.doi.org/10.1590/0102-311X00093214>
13. Gomes RRFM, Batista JR, Ceccato MGB, Kerr LRF, Guimarães MDC. HIV/AIDS knowledge among men who have sex with men: applying the item response theory. *Rev Saúde Pública*. 2014 Apr;48(2):206-15. Doi: <https://doi.org/10.1590/S0034-8910.2014048004911>
14. Souza MF, Santos AFD, Reis IA, Santos MADC, Jorge AO, Machado ATGDM, et al. Care coordination in PMAQ-AB: an item response theory-based analysis. *Rev Saúde Pública*. 2017;51:87. Doi: [10.11606/S1518-8787.2017051007024](https://doi.org/10.11606/S1518-8787.2017051007024)
15. Ensslin L, Ensslin SR, Pacheco GC. A study about safety in football stadiums based on bibliometric analysis of international literature. *Perspect Ciênc Inf*. 2012 Apr/June;17(2):71-91. Doi: <http://dx.doi.org/10.1590/S1413-99362012000200006>.
16. Lacerda RT, Ensslin L, Ensslin SR. A bibliometric analysis of strategy and performance measurement. *Gest Prod*. 2012;19(1):59-78. Doi: <http://dx.doi.org/10.1590/S0104-530X2012000100005>
17. Lee SY, Bender DE, Ruiz RE, Cho YI. Development of an Easy-to-Use Spanish Health Literacy Test. *Health Serv Res*. 2006 Aug;41(4 pt 1):1392-412. Doi: [10.1111/j.1475-6773.2006.00532.x](https://doi.org/10.1111/j.1475-6773.2006.00532.x)
18. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health*. 2013 July;13(1):658. Doi: [10.1186/1471-2458-13-658](https://doi.org/10.1186/1471-2458-13-658)
19. Cai L. A Two-Tier Full-Information Item Factor Analysis Model with Applications. *Psychometrika*. 2010 Dec;75(4):581-612. Doi: [10.1007/S11336-010-9178-0](https://doi.org/10.1007/S11336-010-9178-0)
20. Takegami M, Suzukamo Y, Wakita T, Noguchi H, Chin K, Kadotani H, et al. Development of a Japanese version of the Epworth Sleepiness Scale (JESS) based on item response theory. *Sleep Med*. 2009 May;10(5):556-65. Doi: [10.1016/j.sleep.2008.04.015](https://doi.org/10.1016/j.sleep.2008.04.015)
21. Ravens-Sieberer U, Herdman M, Devine J, Otto C, Bullinger M, Rose M, et al. The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. *Qual Life Res*. 2014 Apr;23(3):791-803. Doi: [10.1007/s11136-013-0428-3](https://doi.org/10.1007/s11136-013-0428-3)
22. Lee SYD, Stucky B, Lee J, Rozier R, Bender D. Short assessment of health literacy-spanish and english: a comparable test of health literacy for spanish and english speakers. *Health Serv Res*. 2010 Aug;45(4):1105-20. Doi: [10.1111/j.1475-6773.2010.01119.x](https://doi.org/10.1111/j.1475-6773.2010.01119.x)
23. Stochl J, Jones PB, Croudace TJ. Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. *BMC Med Res Methodol*. 2012 June;12(1):74. Doi: [10.1186/1471-2288-12-74](https://doi.org/10.1186/1471-2288-12-74)
24. Stansbury JP, Ried LD, Velozo CA. Unidimensionality and Bandwidth in the Center for Epidemiologic Studies Depression (CES-D) Scale. *J Pers Assess*. 2006 Feb;76(1):10-22. Doi: [10.1207/s15327752jpa8601\\_03](https://doi.org/10.1207/s15327752jpa8601_03)
25. Boyer L, Simeoni MC, Loundou A, D'Amato T, Reine G, Lancon C, et al. The development of the S-QoL 18: A shortened quality of life questionnaire for patients with schizophrenia. *Schizophr Res*. 2010 Aug;121(1-3):241-50. Doi: [10.1016/j.schres.2010.05.019](https://doi.org/10.1016/j.schres.2010.05.019)
26. Jenkinson C, Fitzpatrick R, Garratt A, Peto V, Stewart-Brown S. Can item response theory reduce patient burden when measuring health status in neurological disorders? Results from Rasch analysis of the SF-36 physical functioning scale (PF-10). *J Neurol Neurosurg Psychiatry*. 2001 Aug;71(2):220-4. Doi: [10.1136/jnnp.71.2.220](https://doi.org/10.1136/jnnp.71.2.220)
27. Gulliford MC, Mahabir D, Rocke B. Reliability and validity of a short form household food security scale in a Caribbean community. *BMC Public Health*. 2004 June;4:22. Doi: [10.1186/1471-2458-4-22](https://doi.org/10.1186/1471-2458-4-22)
28. Metz SM, Wyrwich KW, Babu AN, Kroenke K, Tierney WM, Wolinsky FD. A comparison of traditional and Rasch cut points for assessing clinically important change in health-related

quality of life among patients with asthma. *Qual Life Res.* 2006 Dec;15(10):1639-49. Doi: [10.1007/s11136-006-0036-6](https://doi.org/10.1007/s11136-006-0036-6)

29. Filiatrault J, Gauvin L, Fournier M, Parisien M, Robitaille Y, Laforest S, et al. Evidence of the psychometric qualities of a simplified version of the activities-specific balance confidence scale for community-dwelling seniors. *Arch Phys Med Rehabil.* 2007 May;88(5):664-72. Doi: [10.1016/j.apmr.2007.02.003](https://doi.org/10.1016/j.apmr.2007.02.003)

30. Jette AM, Haley SM, Ni P, Olarsch S, Moed R. Creating a computer adaptive test version of the late-life function and disability instrument. *J Gerontol A Biol Sci Med Sci.* 2008 Nov;63(11):1246-56. PMID: [19038841](https://pubmed.ncbi.nlm.nih.gov/19038841/)

31. Tao W, Haley SM, Coster WJ, Ni P, Jette AM. An exploratory analysis of functional staging using an item response theory approach. *Arch Phys Med Rehabil.* 2008 June;89(6):1046-53. Doi: [10.1016/j.apmr.2007.11.036](https://doi.org/10.1016/j.apmr.2007.11.036)

32. Allen DD, Gorton GE, Oeffinger DJ, Tylkowski C, Tucker CA, Haley SM. Analysis of the pediatric outcomes data collection instrument in ambulatory children with cerebral palsy using confirmatory factor analysis and item response theory methods. *J Pediatr Orthop.* 2008 Mar;28(2):192-8. Doi: [10.1097/BPO.0b013e3181652185](https://doi.org/10.1097/BPO.0b013e3181652185)

33. Forjaz MJ, Ayala A, Rodriguez-Blazquez C, Frades-Payo B, Martinez-Martin P, Longitudinal Parkinson's Disease Patient Study, et al.. Assessing autonomic symptoms of parkinson's disease with the scopa-aut: a new perspective from rasch analysis. *Eur J Neurol.* 2010 Feb;17(2):273-9. Doi: [10.1111/j.1468-1331.2009.02835.x](https://doi.org/10.1111/j.1468-1331.2009.02835.x)

34. Forjaz MJ, Rodriguez-Blázquez C, Martinez-Martin P, Longitudinal Parkinson's Disease Patient Study Group. Rasch analysis of the hospital anxiety and depression scale in Parkinson's disease. *Mov Disord.* 2009;24(4):526-32. Doi: [10.1002/mds.22409](https://doi.org/10.1002/mds.22409)

35. Haley SM, Fragala-Pinkham MA, Dumas HM, Ni P, Gorton GE, Watson K, et al. Evaluation of an item bank for a computerized adaptive test of activity in children with cerebral palsy. *Phys Ther.* 2009;89(6):589-600. Doi: [10.2522/ptj.20090007](https://doi.org/10.2522/ptj.20090007).

36. Haley SM, Ni P, Jette AM, Tao W, Moed R, Meyers D, et al. Replenishing a computerized adaptive test of patient-reported daily activity functioning. *Qual Life Res.* 2009;18(4):461-71. Doi: [10.1007/s11136-009-9463-5](https://doi.org/10.1007/s11136-009-9463-5)

37. Jette AM, McDonough CM, Ni P, Haley SM, Hambleton RK, Olarsch S, et al. A functional difficulty and functional pain instrument for

hip and knee osteoarthritis. *Arthritis Res Ther.* 2009 July;11(4):107-18. Doi: [10.1186/ar2760](https://doi.org/10.1186/ar2760)

38. Brodersen J, Thorsen H, Kreiner S. Consequences of screening in lung cancer: development and dimensionality of a questionnaire. *Value Health.* 2010 Aug;13(5):601-12. Doi: [10.1111/j.1524-4733.2010.00697.x](https://doi.org/10.1111/j.1524-4733.2010.00697.x)

39. Cherepanov D, Palta M, Fryback DG. Underlying dimensions of the five health-related quality-of-life measures used in utility assessment. *Med Care.* 2010 Aug;48(8):718-25. Doi: [10.1097/MLR.0b013e3181e35871](https://doi.org/10.1097/MLR.0b013e3181e35871)

40. Conrad K, Riley B, Conrad K, Chan Y, Dennis M. Validation of the Crime and Violence Scale (CVS) against the rasch measurement model including differences by gender, race, and age. *Eval Rev.* 2010 Apr;34(2):83-115. Doi: [10.1177/0193841X10362162](https://doi.org/10.1177/0193841X10362162)

41. Perron BE, Vaughn MG, Howard MO, Bohnert A, Guerrero E. Item response theory analysis of DMS-IV criteria for inhalant-use disorders in adolescents. *J Stud Alcohol Drugs.* 2010 July;71(4):607-14. PMID: 20553671

42. Villaveces A, Stucky BD, Runyan CW, Moore KS, Berlin SP. The development of an instrument for evaluating core competencies in violence and injury prevention. *J Public Health Manag Pract.* 2010 July/Aug;16(4):337-44. Doi: [10.1097/PHH.0b013e3181bede2a](https://doi.org/10.1097/PHH.0b013e3181bede2a)

43. Conrad KJ, Iris M, Ridings JW, Langley K, Anetzberger GJ. Self-report measure of psychological abuse of older adults. *Gerontologist.* 2011 June;51(3):354-66. Doi: [10.1093/geront/gnq103](https://doi.org/10.1093/geront/gnq103)

44. Dregan A, Grieve A, Van Staa T, Gulliford MC. Potential application of item-response theory to interpretation of medical codes in electronic patient records. *BMC Med Res Methodol.* 2011 Dec;11(1):168. Doi: <https://doi.org/10.1186/1471-2288-11-168>

45. Fang J, Fleck MP, Green A, McVilly K, Hao Y, Tan W, et al. The response scale for the intellectual disability module of the WHOQOL: 5-point or 3-point? *J Intellect Disabil Res.* 2011 June;55(6):537-49. Doi: [10.1111/j.1365-2788.2011.01401.x](https://doi.org/10.1111/j.1365-2788.2011.01401.x)

46. Riley WT, Pilkonis P, Cella D. Application of the National Institutes of Health patient-reported outcomes measurement information system (PROMIS®) to mental health research. *J Ment Health Policy Econ.* 2011 Dec;14(4):201-8. PMID: 22345362

47. Bann CM, Kobau R, Lewis MA, Zack MM, Luncheon C, Thompson WW. Development and psychometric evaluation of the public health surveillance well-being scale. *Qual Life Res.*

- 2012 Aug;21(6):1031-43. Doi: [10.1007/s11136-011-0002-9](https://doi.org/10.1007/s11136-011-0002-9)
48. Chan KS, Kasper JD, Brandt J, Pezzin LE. Measurement equivalence in ADL and IADL difficulty across international surveys of aging: findings from the HRS, SHARE, and ELSA. *J Gerontol B Psychol Sci Soc Sci*. 2012 Jan;67B(1):121-32. Doi: [10.1093/geronb/gbr133](https://doi.org/10.1093/geronb/gbr133)
49. Fang J, Power M, Lin Y, Zhang J, Hao Y, Chatterji S. Development of short versions for the WHOQOL-OLD Module. *Gerontologist*. 2012 Feb;52(1):66-78. Doi: [10.1093/geront/gnr085](https://doi.org/10.1093/geront/gnr085)
50. Jette AM, Tulsy DS, Ni P, Kisala PA, Slavin MD, Dijkers MP, et al. Development and initial evaluation of the spinal cord injury-functional index. *Arch Phys Med Rehabil*. 2012 Oct; 93(10):1733-50. Doi: [10.1016/j.apmr.2012.05.008](https://doi.org/10.1016/j.apmr.2012.05.008)
51. Lee J, Stucky B, Rozier G, Lee SY, Zeldin LP. Oral health literacy assessment: development of an oral health literacy instrument for spanish speakers. *J Public Health Dent*. 2013;73(1):1-8. Doi: [10.1111/jphd.12000](https://doi.org/10.1111/jphd.12000)
52. Marfeo EE, Ni P, Haley SM, Bogusz K, Meterko M, McDonough CM, et al. Scale refinement and initial evaluation of a behavioral health function measurement tool for work disability evaluation. *Arch Phys Med Rehabil*. 2013 Sept;94(9):1679-86. Doi: [10.1016/j.apmr.2013.03.012](https://doi.org/10.1016/j.apmr.2013.03.012)
53. Marfeo EE, Ni P, Haley SM, Jette AM, Bugosz K, Meterko M, et al. Development of an instrument to measure behavioral health function for work disability: item pool construction and factor analysis. *Arch Phys Med Rehabil*. 2013 Sept;94(9):1670-8. Doi: [10.1016/j.apmr.2013.03.013](https://doi.org/10.1016/j.apmr.2013.03.013)
54. McDonough CM, Jette AM, Ni P, Bogusz K, Marfeo EE, Brandt DE, et al. Development of a self-report physical function instrument for disability assessment: item pool construction and factor analysis. *Arch Phys Med Rehabil*. 2013 Sept;94(9):1653-60. Doi: [10.1016/j.apmr.2013.03.011](https://doi.org/10.1016/j.apmr.2013.03.011)
55. Parsons JT, Rendina HJ, Ventuneac A, Cook KF, Grov C, Mustanski B. A psychometric investigation of the hypersexual disorder screening inventory among highly sexually active gay and bisexual men: an item response theory analysis. *J Sex Med*. 2013 Dec;10(12):3088-101. Doi: [10.1111/jsm.12117](https://doi.org/10.1111/jsm.12117)
56. Waller J, Ostini R, Marlow LAV, McCaffery K, Zimet G. Validation of a measure of knowledge about human papillomavirus (HPV) using item response theory and classical test theory. *Prev Med*. 2013 Jan;56(1):35-40. Doi: <https://doi.org/10.1016/j.yjmed.2012.10.028>
57. Bevans M, Ross A, Cella D. Patient-Reported Outcomes Measurement Information System (PROMIS): Efficient, standardized tools to measure self-reported health and quality of life. *Nurs Outlook*. 2014 Sept/Oct;62(5):339-45. Doi: [10.1016/j.outlook.2014.05.009](https://doi.org/10.1016/j.outlook.2014.05.009)
58. Hahn EA, DeWalt DA, Bode RK, Garcia SF, DeVellis RF, Correia H, et al. New English and Spanish social health measures will facilitate evaluating health determinants. *Health Psychol*. 2014 May;33(5):490-9. Doi: [10.1037/hea0000055](https://doi.org/10.1037/hea0000055)
59. Onishi Y, Wakita T, Fukuhara S, Noguchi Y, Okada M, Sakaida I, et al. Development and validation of a symptom scale specific for ascites accompanied with cirrhosis: the ASI-7. *Clin Transl Gastroenterol*. 2014 Feb;5(2):e48. Doi: [10.1038/ctg.2013.20](https://doi.org/10.1038/ctg.2013.20)
60. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware Jr JE. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *J Clin Epidemiol*. 2014 May;67(5):516-26. Doi: [10.1016/j.jclinepi.2013.10.024](https://doi.org/10.1016/j.jclinepi.2013.10.024)
61. Spiegel B, Hays R, Bolus R, Melmed G, Chang L, Whitman C et al. Development of the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) Gastrointestinal Symptom Scales. *Am J Gastroenterol*. 2014 Nov;109(11):1804-14. Doi: [10.1038/ajg.2014.237](https://doi.org/10.1038/ajg.2014.237)
62. Zhong Q, Gelaye B, Fann JR, Sanchez SE, Williams MA. Cross-cultural validity of the spanish version of PHQ-9 among pregnant peruvian women: a raschitem response theory analysis. *J Affect Disord*. 2014;158:148-53. Doi: [10.1016/j.jad.2014.02.012](https://doi.org/10.1016/j.jad.2014.02.012)
63. Chan KS, Gross AL, Pezzin LE, Brandt J, Kasper JD. Harmonizing measures of cognitive performance across international surveys of aging using item response theory. *J Aging Health*. 2015 Dec;27(8):1392-414. Doi: [10.1177/0898264315583054](https://doi.org/10.1177/0898264315583054)
64. Epstein J, Osborne RH, Elsworth GR, Beaton DE, Guillemin F. Cross-cultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. *J Clin Epidemiol*. 2015 Apr;68(4):360-9. Doi: [10.1016/j.jclinepi.2013.07.013](https://doi.org/10.1016/j.jclinepi.2013.07.013)
65. Gerrard P, Kazis LE, Ryan CM, Shie VL, Holavanahalli R, Lee A, et al. Validation of the Community Integration Questionnaire in the adult burn injury population. *Qual Life Res*. 2015 Nov;24(11):2651-5. Doi: [10.1007/s11136-015-0997-4](https://doi.org/10.1007/s11136-015-0997-4)

66. Honda M, Wakita T, Onishi Y, Nunobe S, Miura A, Nishigori T, et al. Development and validation of a disease-specific instrument to measure diet-targeted quality of life for postoperative patients with esophagogastric cancer. *Ann Surg Oncol*. 2015 Dec;22(S3):848-54. Doi: <https://doi.org/10.1245/s10434-015-4696-8>

67. Michel P, Auquier P, Baumstarck K, Pelletier J, Loundou A, Ghattas B, et al. Development of a cross-cultural item bank for measuring quality of life related to mental health in multiple sclerosis patients. *Qual Life Res*. 2015 Sept;24(9):2261-71. Doi: [10.1007/s11136-015-0948-0](https://doi.org/10.1007/s11136-015-0948-0)

68. Schalet BD, Rothrock NE, Hays RD, Kazis LE, Cook KF, Rutsohn JP, et al. Linking Physical and Mental Health Summary Scores from the Veterans RAND 12-Item Health Survey (VR-12) to the PROMIS® Global Health Scale. *J Gen Intern Med*. 2015 Oct;30(10):1524-30. Doi: [10.1007/s11606-015-3453-9](https://doi.org/10.1007/s11606-015-3453-9)

69. Teicher MH, Parigger A. The 'Maltreatment and Abuse Chronology of Exposure' (MACE) Scale for the Retrospective Assessment of Abuse and Neglect During Development. *PLoS One*. 2015 Feb;10(2):e0117423. Doi: [10.1371/journal.pone.0117423](https://doi.org/10.1371/journal.pone.0117423)

70. Coster WJ, Ni P, Slavin MD, Kisala PA, Nandakumar R, Mulcahey MJ, Tulsy DS, Jette AM. Differential item functioning in the Patient Reported Outcomes Measurement Information System Pediatric Short Forms in a sample of children and adolescents with cerebral palsy. *Dev Med Child Neurol*. 2016 Nov;58(11):1132-8. Doi: [10.1111/dmcn.13138](https://doi.org/10.1111/dmcn.13138)

71. McDonough CM, Ni P, Coster WJ, Haley SM, Jette AM. Development of an IRT-Based short form to assess applied cognitive function in outpatient rehabilitation. *Am J Phys Med Rehabil*. 2016 Jan;95(1):62-71. Doi: [10.1097/PHM.0000000000000340](https://doi.org/10.1097/PHM.0000000000000340)

72. Marino M, Soley-Bori M, Jette AM, Slavin MD, Ryan CM, Schneider JC, et al. Measuring the social impact of burns on survivors. *J Burn Care Res*. 2017 Jan/Feb;38(1):377-83. Doi: [10.1097/BCR.0000000000000398](https://doi.org/10.1097/BCR.0000000000000398)

73. Andrade D, Tavares H, Valle R. Teoria da Resposta ao Item: conceitos e aplicações. São Paulo: ABE; 2000.

74. Lord F. Applications of item response theory to practical testing problems. New York: Routledge; 1980.

75. Birnbaum A. Some latent trait models and their use in inferring and examinee's ability. In: Lord FM, Novick MR. Statistical theories of

mental test scores. Menlo Park: Addison-Wesley; 1968.

76. Samejima F. A new family of models for the multiple-choice item. Knoxville: University of Tennessee; 1979.

77. Muraki E. A generalized partial credit model: application of an EM algorithm. *Applied Psychological Measurement*. 1992 June;16(2):159-76. Doi: <https://doi.org/10.1177/014662169201600206>

78. Ferreira EV, Azevedo CLN. Contribuições ao estudo do Modelo de Crédito Parcial Generalizado. *Est aval educ*. 2017 Jan/Apr; 28(67):236-54. Doi: <http://dx.doi.org/10.18222/eae.v0ix.3385>

79. Sisto FF. Differential item functioning. *Psico-USF [Internet]*. 2006 Jan/June [cited 2017 Nov 15];11(1):35-43. Available from: <http://www.scielo.br/pdf/pusf/v11n1/v11n1a05.pdf>

80. Dalcin CB, Backes DS, Dotto JI, Souza MHT, Moreschi C, Büscher A. Social determinants of health that influence the healthy living process in a vulnerable community. *J Nurs UFPE on line*. 2016 June;10(6):1963-70. Doi: 10.5205/reuol.9199-80250-1-SM1006201608

Submissão: 15/02/2018

Aceito: 19/03/2018

Publicado: 01/06/2018

#### Correspondência

Diego Eller Gomes

Rua Professor Milton Sullivan, 28 / Ap. 103

Bairro Carvoeira

CEP: 88040-620 – Florianópolis (SC), Brasil