

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA**



DISSERTAÇÃO DE MESTRADO
TEORIA DA RESPOSTA AO ITEM NA AVALIAÇÃO DE
QUALIDADE DE VIDA DE IDOSOS

Katrine Guewehr

Orientador: Prof. Dr. João Riboldi

Porto Alegre, maio de 2007.

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA**



DISSERTAÇÃO DE MESTRADO
TEORIA DA RESPOSTA AO ITEM NA AVALIAÇÃO DE
QUALIDADE DE VIDA DE IDOSOS

Katrine Guewehr

Orientador: Prof. Dr. João Riboldi

A apresentação desta dissertação é exigência do Programa de Pós-graduação em Epidemiologia, Universidade Federal do Rio Grande do Sul, para obtenção do título de Mestre.

Porto Alegre, Brasil.
2007

BANCA EXAMINADORA

Prof. Dr. João Feliz Duarte de Moraes [Departamento de Estatística], [UFRGS];

Prof. Dra. Maria Inês Azambuja, [Programa de Pós-graduação em Epidemiologia], [UFRGS];

Prof. Dr. Álvaro Vigo, [Programa de Pós-graduação em Epidemiologia], [UFRGS].

AGRADECIMENTOS

Em primeiro lugar gostaria de agradecer ao meu orientador e acima de tudo amigo, Prof. Dr. João Riboldi, pelo constante incentivo, carinho e enriquecimento da minha formação profissional.

Ao Prof. Dr. João Feliz Duarte de Moraes por ceder seus dados e tornar possível a realização deste trabalho.

Ao Prof. Dr. Dalton de Andrade, pelo empréstimo de materiais sobre Teoria de Resposta ao Item, e sugestões importantes para as análises.

Ao suporte financeiro do CNPq, imprescindível nessa etapa.

À Universidade Federal do Rio Grande do Sul e ao Programa de Pós-graduação em Epidemiologia da Faculdade de Medicina.

A todos meus amigos, colegas do PPGEPI, funcionários do Núcleo de Assessoria Estatística e Setor de Recursos Computacionais do Instituto de Matemática da UFRGS, que de alguma forma me ajudaram e apoiaram.

E à minha família, especialmente minha irmã, pela força nas horas mais complicadas.

SUMÁRIO

Abreviaturas e Siglas.....	8
Resumo.....	10
Abstract.....	11
Lista de Quadros.....	12
Lista de Tabelas.....	13
Lista de Figuras.....	15
1. APRESENTAÇÃO.....	16
2. INTRODUÇÃO.....	17
3. REVISÃO DA LITERATURA.....	19
3.1 Técnicas Tradicionais de Análise.....	19
3.1.1 Teoria Clássica dos Testes ou das Medidas.....	19
3.1.2 Análise Fatorial.....	21
3.2 Teoria da Resposta ao Item.....	22
3.2.1 Histórico da Teoria da Resposta ao Item.....	22
3.2.2 Modelos da Resposta ao Item.....	23
3.2.2.1 Modelo Logístico de 3 Parâmetros.....	24
3.2.2.2 Modelo Logístico de 1 e 2 Parâmetros.....	26
3.2.2.3 Outros Modelos de Teoria da Resposta ao Item.....	27
3.2.3 Função de Informação do Item.....	28
3.2.4 Função de Informação do Questionário.....	29
3.2.5 Escala de Habilidade.....	30
3.2.6 Suposições do Modelo.....	30

3.2.7 Métodos de Estimação.....	32
3.2.7.1 Estimação das Habilidades dos Indivíduos.....	33
3.2.7.2 Estimação dos Parâmetros dos Itens.....	34
3.2.7.3 Estimação Conjunta dos Parâmetros dos Itens e das Habilidades dos Indivíduos.....	34
3.2.8 Ajuste do Modelo.....	36
3.3 Avaliação de Qualidade de Vida.....	37
3.3.1 Instrumentos de Avaliação de Qualidade de Vida.....	38
3.3.1.1 Instrumentos de Avaliação de Qualidade de Vida desenvolvidos pela Organização Mundial de Saúde.....	40
3.3.1.2 Escala de Qualidade de Vida de Flanagan.....	44
3.4 Avaliação de Qualidade de Vida em Idosos.....	46
3.5 Teoria da Resposta ao Item na Avaliação de Qualidade de Vida.....	49
4. OBJETIVOS.....	55
5. REFERÊNCIAS BIBLIOGRÁFICAS.....	56
6. ARTIGO.....	64
7. CONCLUSÕES E CONSIDERAÇÕES FINAIS.....	113
8. ANEXOS	
Anexo A - Projeto de Pesquisa.....	119
Anexo B - Aprovação pela Comissão de Pesquisa.....	131
Anexo C – Autorização do responsável para utilizar a base de dados.....	133
Anexo D – Escala de Qualidade de Vida de Flanagan.....	135

Anexo E – Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS (WHOQOL-100).....	137
Anexo F – Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS – Versão abreviada (WHOQOL-BREVE).....	142
Anexo G – Estimativas das habilidades (traços latentes) dos idosos.....	145
Anexo H – Tabelas e adicionais que não foram colocadas no artigo.....	157
Anexo I – Resultados Adicionais que não foram colocados no artigo.....	161
Anexo J – Estatísticas de Concordância.....	174
Anexo K – Recursos Computacionais.....	177

ABREVIATURAS E SIGLAS

ADL: *Activities of dayling general health questionnaire*
 AF: Análise Fatorial
 ALDS: *AMC Linear Disability Score*
 ALSAQ-5: *Amyotrophic Lateral Sclerosis Assesment Questionnaire-40 itens*
 ALSAQ-40: *Amyotrophic Lateral Sclerosis Assesment Questionnaire-5 itens*
 BDI: *Beck Depression Inventory*
 BHS: *Beck Hopelessness Scale*
 CAT: *Computer adapative testing*
 CCI: Curva característica do item
 DIF: *Differential item functioning*
 DLTV: *Daily Living Tasks Dependent in Vision*
 EAP: Esperança a posteriori
 E.L.: Escore latente
 EM: Esperança e maximização
 EORTC QLQ C30: *European Organization for Research and Treatment of Cancer Quality of Life Group*
 EP: Erro padrão
 EQVF: Escala de Qualidade de Vida de Flanagan
 ESSI: *earth search science*
 ETS: *Educational Testing Service*
 EUA: Estados Unidos da América
 FACT: *The Functional Assessment of Cancer Therapy*
 FACT-An: *The Functional Assessment of Cancer Therapy-Anemia*
 FACIT-F: *Functional Assessment of Chronic Illness Therapy Fatigue Scale*
 HAQ: *Health assessment questionnaire*
 HC SUS – HIV: *HIV Cost and Services Utilization Study*
 HIT: *Headache Impact Test*
 HIT-6: *Headache Impact Test-6 itens*
 HIV: *Human Immunodeficiency Virus*
 HRQOL: *Health related quality of life*
 IBGE: Instituto Brasileiro de Geografia e Estatística
 IBRASA: Instituição Brasileira de Difusão Cultural S.A.
 ISS: Índice de saúde subjetiva
 MacDQoL: *Macular degeneration on quality of life questionnaire*
 MacDQoL-22: *Macular degeneration on quality of life questionnaire- 22 itens*
 MAP: Máximo a posteriori
 ML-1: Modelo logístico de 1 parâmetro
 ML-2: Modelo logístico de 2 parâmetros
 ML-3: Modelo logístico de 3 parâmetros
 MSQ: *Migraine Specific Quality of Life Questionnaire*
 MV: Máxima verossimilhança
 MVC: Máxima verossimilhança conjunta
 MVM: Máxima verossimilhança marginal
 NHP: *Nottingham Health Profile*
 NSHI: *National Service of Headache Impact*

OMS: Organização Mundial de Saúde
PAR-Q: *Questionnaire relating to readiness for physical activities*
PF-10: *Physical functioning*
PIPP: *Perceived Impact of Problem Profile*
QOL: *Quality of life*
QV: Qualidade de vida
SDS: *Symptom distress scale*
SF-36: *Medical Outcomes Study Short-Form*
TCM: Teoria Clássica das Medidas
TCT: Teoria Clássica dos Testes
TRI: Teoria da Resposta ao Item
UFRGS: Universidade Federal do Rio Grande do Sul
USP: Universidade de São Paulo
WHO: *World Health Organization*
WHOQOL: World Health Organization Quality of Life Group
WHOQOL-BREVE: Instrumento de Avaliação de Qualidade de Vida da OMS -
versão abreviada
WHOQOL-OLD: Instrumento de Avaliação de Qualidade de Vida da OMS para
idosos
WHOQOL-100: Instrumento de Avaliação de Qualidade de Vida da OMS
WOMAC: *Western Ontario MacMaster Questionnaire*

RESUMO

INTRODUÇÃO: A importância da avaliação da qualidade de vida no contexto da saúde vem ganhando crescente valorização tanto no meio acadêmico, quanto na sociedade, e esse fenômeno vem sendo amplamente reconhecido nas últimas décadas. Os idosos não fogem a esse panorama e almejam envelhecer com qualidade de vida. A Teoria Clássica dos Testes (TCT) foi, e ainda é, muito utilizada, mas apresenta várias limitações, e os modelos de Teoria da Resposta ao Item (TRI) lidam com consistência e eficiência na avaliação de tais desfechos, utilizando mais intensamente a informação presente nos instrumentos.

OBJETIVOS: Estruturar a TRI e aplicá-la aos resultados obtidos no uso de três instrumentos de avaliação de qualidade de vida, e comparar estes resultados com os resultados provenientes dos métodos tradicionais de análises, via TCT e Análise Fatorial. Comparar a equivalência do WHOQOL-BREVE e WHOQOL-100.

MÉTODO: Foi utilizado o Modelo Logístico de 2 Parâmetros na estimação da qualidade de vida global e por domínios a partir dos resultados obtidos com os instrumentos de avaliação de qualidade de vida elaborados pela OMS (WHOQOL-100 e WHOQOL-BREVE) e pela Escala de Qualidade de Vida de Flanagan (EQVF) em 400 idosos socialmente ativos da região metropolitana de Porto Alegre, Rio Grande do Sul. As respostas foram dicotomizadas após a realização da pesquisa. Foram verificadas as consistências internas de cada escala através do coeficiente α de Cronbach, e realizadas análises clássicas para cada instrumento, análises fatoriais, análises de variância com os escores brutos e os escores latentes obtidos na TRI, análises de correlação linear de Pearson entre os escores e a verificação da concordância através do coeficiente Kappa a partir da classificação com base na mediana dos escores.

RESULTADOS: As três escalas apresentaram valores altos de fidedignidade. Com base na aplicação do modelo de TRI na EQVF, verificou-se que os idosos têm maior dificuldade de sentirem-se satisfeitos com sua aprendizagem, e que indivíduos com um bom nível de autoconhecimento, trabalho, boa comunicação e oportunidades de lazer tendem a ter um nível superior de qualidade de vida. Os escores brutos e latentes apresentaram uma correlação de 0,972. Os fatores significantes na análise de variância para os escores brutos foram saúde percebida, níveis de depressão e desesperança. E com os escores latentes, apenas saúde percebida e depressão foram significantes a 1 e 5%. Na EQVF, os idosos apresentaram concordância de classificação da qualidade de vida, e o item relativo à vida íntima foi retirado das análises devido à inconsistência. Através do WHOQOL-BREVE, verificou-se que estes idosos têm maior dificuldade de estarem satisfeitos com sua qualidade de vida em geral, e dão grande importância para a capacidade de desempenhar as atividades cotidianas. Os itens referentes à vida sexual, dinheiro suficiente para satisfazer suas necessidades e satisfação com o acesso aos serviços de saúde não se mostraram válidos para estimar a qualidade de vida de idosos na TRI, sendo retirados da estimação das habilidades. Os fatores influentes nos escores brutos e latentes são os mesmos: saúde percebida, níveis de depressão e desesperança, para o WHOQOL-BREVE e WHOQOL-100. As correlações entre os escores brutos e latentes para o WHOQOL-BREVE e WHOQOL-100 são 0,968 e 0,940, respectivamente. Foi encontrada concordância na classificação dos escores de qualidade de vida para o WHOQOL-100 e WHOQOL-BREVE. Com base na TRI foram encontrados baixos níveis de satisfação com a memória, vida sexual e situação financeira no WHOQOL-100. O instrumento apresentou 35 itens inconsistentes.

CONCLUSÕES: Os resultados via TRI não diferem significativamente dos resultados dos métodos tradicionais, mas confirmam a TRI como um método promissor para avaliação de qualidade de vida.

Descritores: qualidade de vida, escalas, idoso, análise fatorial.

ABSTRACT

INTRODUCTION: The life's quality evaluation importance on the health's context has been valorized even on the academic, or on the society way and this phenomenon is very recognized last decades. The elderly aren't out of this panorama and they long for to ageing with life's quality. The Classical Test Theory (CTT) was and is still used, buy it shows several limitations otherwise the Item Response Theory models (IRT) deals with the consistence and efficiency in this outcome evaluation, using more intensely the presented information on the instruments. **OBJECTIVES:** Structuralize the IRT and apply the obtained results on the three life's quality evaluation's instruments, comparing these with the results of traditional methods by CTT and Factor Analysis, and compare the equivalence between WHOQOL-BREF and WHOQOL-100. **METHODS:** It used two parameters logistic model in the global life's quality estimation and with domains come from the results obtained to the life's quality evaluation instruments elaborated for WHO (WHOQOL-100 and WHOQOL-BREF) and for Flanagan Quality of Life Scale (FQLS) in 400 elderly socially active in metropolitan region of Porto Alegre, RS. The answers were dichotomized after the research being done. It verified the internal consistency for the three scales through reliability analysis (α Cronbach), and realized classical analyses for each instrument, factor analysis, analysis of variance with the brutes scores and latent scores by IRT, Pearson's linear correlation among the scores and the Kappa concordance's statistic through the classification based in the scores median. **RESULTS:** The three scales presented high level of reliability. Based on the model's application of IRT in the FQLS, it verified that elderly have more difficulties on felling satisfaction with their learning, and that good levels of self-knowledge, job, good communication and leisure opportunities are indicators of a better life's quality. The scores brute and latent presented a correlation of 0,972. The significant variables in the brute's analysis of variance were perceived health, depression level and hopelessness level. And with latent scores only the variables perceived health and depression level were significant at 1 and 5%. In FQLS, the elderly did get agreement in the life's quality classification, and the sexual life item was out from analysis because of inconsistency. Through the WHOQOL-BREF, it verified that these elderly have more difficulties of being satisfied with life's quality in general and gave a lot of importance to the capacity to play their daily activities. The items for sexual life, money enough to satisfy their necessities and satisfaction with the service health access didn't show valid to estimate the life's quality in elderly on TRI, being removed of the ability's estimation. The influent factors in both scores were the same: perceived health, depression and hopelessness levels to WHOQOL-BREF and WHOQOL-100. The correlations between brute and latent scores to WHOQOL-BREF and WHOQOL-100 were 0,968 and 0,940, respectively. It found concordance in life's quality scores to the WHOQOL-BREF and WHOQOL-100. Based on IRT was founded low level of satisfaction with memory, sexual life and financial situation on WHOQOL-100. The instrument presented 35 inconsistent items. **CONCLUSIONS:** The results by IRT didn't change that much from the traditional methods, but IRT was confirmed as a promising to evaluation of life's quality.

Describers: life's quality, scales, elderly, factor analysis.

LISTA DE QUADROS

QUADRO 1 – Modelos Unidimensionais de Teoria de Resposta ao Item.....	28
QUADRO 2 – Estágios do desenvolvimento do WHOQOL-100.....	40
QUADRO 3 – Domínios e facetas do WHOQOL-100.....	42
QUADRO 4 – Domínios e facetas do WHOQOL-BREVE.....	44
QUADRO 5 – Dimensões/Itens da Escala de Qualidade de Vida de Flanagan.....	45

LISTA DE TABELAS

Artigo

TABELA 1 – Descrição dos dados dos 400 idosos socialmente ativos da região metropolitana de Porto Alegre.....	98
TABELA 2 – Distribuição dos itens da Escala de Qualidade de Vida de Flanagan, estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses....	99
TABELA 3 – Estatísticas clássicas para os itens da EQVF.....	100
TABELA 4 – Distribuição dos itens do Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS – versão abreviada (WHOQOL-BREVE), estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses.....	101
TABELA 5 – Estatísticas clássicas para os itens do WHOQOL-BREVE.....	103
TABELA 6 – Matriz das correlações lineares de Pearson entre os escores latentes dos domínios e dos escores latentes globais (WHOQOL-BREVE).....	104
TABELA 7 – Distribuição dos itens do Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS (WHOQOL-100), estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses.....	105
TABELA 8 – Estatísticas clássicas para os itens do WHOQOL-100.....	109
TABELA 9 – Matriz das correlações lineares de Pearson entre os escores latentes dos domínios e dos escores latentes globais (WHOQOL-100).....	112

Anexos

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens.....	145
TABELA 2 – Análise Fatorial para a Escala de Qualidade de Vida de Flanagan (EQVF), sem item 5.....	157
TABELA 3 – Análise Fatorial para WHOQOL – BREVE, sem os itens F153, F181, F193.....	158
TABELA 4 – Análise Fatorial para WHOQOL-100 sem os 35 itens inconsistentes.....	159

TABELA 5 – Correlações entre os escores latentes (E.L) globais (G) e dos domínios (D) entre os instrumentos WHOQOL-BREVE (WB) e WHOQOL-100 (W100), com a retirada com a retirada dos itens inconsistentes.....	160
--	-----

LISTA DE FIGURAS

Revisão da Literatura

FIGURA 1 – Curva Característica do Item para Modelo Logístico de 3 Parâmetros.....	26
--	----

Artigo

FIGURA 1 – Curva Característica do Item para Modelo Logístico de 2 Parâmetros.....	93
FIGURA 2 – Curvas características dos itens da EQVF.....	94
FIGURA 3 – Curvas características dos itens do WHOQOL-BREVE.....	95
FIGURA 4 – Curvas características dos itens do WHOQOL-100.....	96
FIGURA 5 – Distribuição dos escores brutos e latentes para a EQVF, WHOQOL-BREVE e WHOQOL-100.....	97

1. APRESENTAÇÃO

Este trabalho consiste na dissertação de mestrado intitulada “TEORIA DA RESPOSTA AO ITEM NA AVALIAÇÃO DE QUALIDADE DE VIDA DE IDOSOS”, apresentada ao Programa de Pós-Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul, em 30 de maio de 2007. O trabalho é apresentado em três partes, na ordem que segue:

1. Introdução, Revisão da Literatura e Objetivos
2. Artigo
3. Conclusões e Considerações Finais.

Documentos de apoio estão apresentados nos anexos: Projeto de Pesquisa (Anexo A), aprovação pela Comissão de Pesquisa (Anexo B), autorização do responsável para utilizar a base de dados (Anexo C), Escala de Qualidade de Vida de Flanagan (Anexo D), Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS (Anexo E), Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS na versão abreviada (Anexo F), estimativas das habilidades dos idosos (Anexo G), tabelas que não foram colocadas no artigo (Anexo H), resultados adicionais que não foram colocados no artigo (Anexo I), estatísticas de concordância (Anexo J) e recursos computacionais (Anexo K).⁵

2. INTRODUÇÃO

A importância da qualidade de vida no contexto da saúde vem ganhando crescente valorização tanto no meio acadêmico, quanto na sociedade, e esse fenômeno vem sendo amplamente reconhecido nas últimas décadas. Os idosos, objeto de estudo deste trabalho, não fogem à regra, almejando envelhecer com qualidade de vida.

A saúde não é a mera ausência de doença, mas sim um estado de completo bem-estar físico, mental e social, de acordo com a Organização Mundial da Saúde (OMS). Já, o modelo biomédico tradicional, embasado nos desfechos clínicos, bioquímicos, fisiológicos e moleculares, encontra-se carente de informações sobre o bem estar psicológico (mental) e social dos pacientes, além de sua percepção sobre a própria saúde.

Visando preencher essas lacunas, surgiram os instrumentos de avaliação de qualidade de vida, multidimensionais por natureza, que levam em consideração múltiplos aspectos da vida dos indivíduos (físico, psicológico, social, ambiental, espiritual...) avaliados pelo próprio sujeito. O essencial está nessa avaliação subjetiva/percebida pelo indivíduo, pois os tratamentos clínicos devem visar agregar anos com qualidade de vida aos pacientes, resultando assim numa medida importante de desfecho.

Os testes psicológicos que avaliam principalmente qualidade de vida têm hoje uma inegável importância, tornando-se indispensáveis ao nível da prática e da investigação no domínio específico da Psicologia e Saúde, além de outras áreas afins. A sua crescente valorização e utilização trouxeram consigo também acrescidas exigências de validade e de fidedignidade.

Há muitas décadas que a Teoria Clássica dos Testes (TCT) tem sido útil para o desenvolvimento dos testes psicológicos, embora apresente várias limitações, como ser dependente do conjunto de itens que compõem o instrumento de medida como um todo, limitando-se assim a sua aplicabilidade (Vianna 1978; Andrade et al. 2000). Na Teoria da Resposta ao Item (TRI), o procedimento de medida utilizado parte do pressuposto de que existe no indivíduo um traço latente (característica individual determinante de como responder itens de um teste), que possui uma relação

probabilística com cada um dos itens, considerando-se ainda que os parâmetros de cada item não dependem dos outros itens do teste. Com esta técnica, é possível verificar e classificar os indivíduos quanto à alguma característica de interesse, como por exemplo, na área educacional, se um item é mais ou menos fácil/difícil, e assim propor novos métodos de ensino ou políticas de aprendizagem. Na TRI, tanto os itens quanto as pessoas são colocados na mesma escala, podendo ser diretamente comparáveis, o que é uma grande vantagem sobre os outros procedimentos.

A TRI não entra em contradição com os métodos clássicos, mas sim, traz uma nova proposta de análise estatística, centrada nos itens, e que preenche algumas lacunas da TCT, e adicionalmente apresenta novos recursos tecnológicos para a avaliação psicológica e educacional (Vendramini et al. 2004), como a criação e utilização de bancos de itens (Bjorner et al. 2003), testes adaptativos computadorizados (Revicki e Cella 1997) tão importantes na avaliação da saúde, o estudo de viés dos itens e a análise de comportamento diferencial entre grupos (Crane et al. 2006).

Este trabalho ao estruturar metodologicamente a Teoria da Resposta ao Item visa ressaltar as potenciais vantagens do uso de modelos de TRI sobre os métodos tradicionais de análise, e assim contribuir para aumentar a eficiência na avaliação dos fenômenos relacionados à qualidade de vida.

3. REVISÃO DA LITERATURA

3.1 Técnicas tradicionais de análise

3.1.1 Teoria Clássica dos Testes ou das Medidas

Na Teoria Clássica dos Testes (TCT), também conhecida como Teoria Clássica das Medidas (TCM), são avaliadas as propriedades psicométricas dos itens (questões), especificamente, a confiabilidade (consistência interna da escala), a análise do item e a validade do construto.

Na TCM, o escore observado num teste é composto de um escore verdadeiro mais o erro de medida, este último assumido como aleatório. A equação básica conhecida como modelo clássico de medida, descreve a relação entre os escores observados, os escores verdadeiros e o erro:

$$X_{ij} = T_i + e_{ij} \quad (1)$$

onde:

X_{ij} é o escore observado para o indivíduo i no teste j ,

T_i é o escore verdadeiro para o indivíduo i no teste j ,

e_{ij} é o erro de medida para o indivíduo i no teste j .

Assume-se que $E(e_{ij}) = 0$, $\text{cov}(T_i, e_{ij}) = 0$ e $\text{cov}(T_i, e_{kj}) = 0$ para $k \neq i$.

O subscrito ou índice j pode indexar formas de testes paralelos ou tempos de administração do mesmo teste. Duas formas são ditas paralelas se elas possuem escores verdadeiros idênticos e se seus erros têm variâncias iguais.

Desde que $\text{cov}(T_i, e_{kj}) = 0$,

$$\text{var}(X_{ij}) = \text{var}(T_i) + \text{var}(e_{ij}) \quad (2)$$

ou, em outra notação,

$$\sigma_X^2 = \sigma_T^2 + \sigma_e^2 \quad (3)$$

A fidedignidade de um teste pode ser definida por três modos, que são equivalentes no modelo clássico: pela razão da variância dos escores verdadeiros com a variância dos escores observados, pela correlação dos escores observados nas duas formas paralelas de um teste ou pelo quadrado da correlação entre os escores observados e os verdadeiros.

A TCM quando comparada à TRI apresenta uma série de limitações teóricas, como:

(i) Os parâmetros clássicos dos itens (dificuldade e discriminação) dependem diretamente da amostra selecionada, e se a mesma não for representativa da população, os parâmetros dos itens não podem ser considerados válidos para esta população.

(ii) A avaliação de aptidões ou habilidades também depende do teste utilizado. Assim, testes diferentes que medem a mesma aptidão irão produzir escores diferentes da mesma aptidão para sujeitos idênticos. O mesmo acontece com testes com índices de dificuldade diferentes.

(iii) A definição do conceito de fidedignidade ou precisão na teoria clássica dos testes constitui também uma fonte de dificuldades. A fidedignidade é a correlação entre escores obtidos de formas paralelas de um teste ou, mais genericamente, como o oposto do erro de medida. Ambos os conceitos apresentam problemas. Primeiramente, é praticamente impossível satisfazer as condições de definição de formas paralelas e, no caso do erro de medida, é postulado que este seja idêntico em todos os examinados, o que é improvável.

(iv) Outro problema da teoria clássica é sua orientação para o teste total e não para o item individual. Toda a informação do item deriva de considerações do teste geral, não se podendo assim determinar como o examinado se comportaria diante de cada item individual, como acontece na TRI, onde o importante é o item e não o escore total da prova ou teste.

(v) Os métodos clássicos de análise necessitam de dados completos ou técnicas de imputação para dados faltantes, ao passo que na TRI isso não é necessário.

Maiores detalhes sobre TCT podem ser encontrados em Vianna (1978), van Belle et al. (2000), Embretson e Reise (2000), Bechger et al. (2003) e Holland e Hoskens (2003).

3.1.2 Análise Fatorial

Os modelos de variáveis latentes são desenvolvidos para investigar a existência e medir estruturas latentes subjacentes a um conjunto de variáveis observáveis, tais como itens em um questionário. São exemplos de modelos de variáveis latentes o modelo de análise fatorial e os modelos de Teoria da Resposta ao Item, dentre outros. O modelo de variáveis latentes mais conhecido é o modelo de análise fatorial (AF), no qual ambas as variáveis, latentes e observadas, são contínuas. Esses modelos buscam explicar o comportamento das variáveis observadas em relação ao comportamento de um conjunto de variáveis não observadas (variáveis latentes ou fatores).

Na literatura relacionando modelos de variáveis latentes, dentre eles, os modelos de TRI e AF, têm-se a equivalência dos modelos de função resposta (*response function approach*) e do modelo de variável subjacente (*underlying function approach*) para dados binários. A função resposta para dados binários consiste na probabilidade de resposta positiva a um item ou de uma forma mais geral, pela curva característica do item. Já a variável subjacente representa o traço latente que supostamente deve estar sendo medido a partir dos itens. Em suma, os parâmetros estimados pelo modelo de variável latente equivalem às cargas fatoriais da AF para dados binários. Para respostas politômicas, não há equivalência entre os dois modelos. Os métodos para dados politômicos são mais recentes e menos estruturados que os métodos para respostas binárias, fazendo-se necessária mais pesquisa e aprofundamento nessa área.

Detalhes sobre os modelos de função resposta e modelo de variável latente para dados politômicos e binário encontram-se em Bartholomew (1987), Albanese (1990) e Takane e Leeuw (1987).

3.2 Teoria da Resposta ao Item

Este capítulo está densamente baseado em Andrade et al. (2000), Embretson e Reise (2000) e Baker (2001).

3.2.1 Histórico da Teoria da Resposta ao Item

Os primeiros modelos de resposta ao item surgiram na década na 50, mas somente nos últimos 15 anos é que vem sendo utilizados em larga escala, devido à sua complexidade e dificuldade computacional e à ausência de programas computacionais eficientes disponíveis. A principal aplicação da Teoria da Resposta ao Item (TRI) ainda vem sendo as avaliações educacionais (Vendramini et al. 2004), mas há uma crescente difusão desta técnica entre as outras áreas de conhecimento (Hambleton 2000), com destaque para a Psicologia (Pasquali e Primi 2003) e Medicina (Cella et al. 2007; Hays e Lipscomb 2007; McHorney e Cohen 2000), além das áreas de marketing, produção de índices sócio-econômicos (Soares 2005) e na gestão pela qualidade total (Alexandre et al. 2002).

Os primeiros modelos consideravam apenas uma habilidade (traço latente), que é uma característica não observada diretamente, de apenas um grupo ou amostra. Foram inicialmente desenvolvidos na forma de uma função ogiva normal e, depois foram descritos para a forma logística, matematicamente mais conveniente.

Lord, estatístico do Educational Testing Service (ETS) dos Estados Unidos, em 1952, foi o primeiro a desenvolver o modelo unidimensional de dois parâmetros (dificuldade e discriminação) baseado na distribuição normal acumulada (ogiva normal), incorporando mais tarde um parâmetro que tratava do problema do acerto casual, surgindo o modelo de três parâmetros. Birnbaum, em 1968, substituiu em ambos os modelos propostos por Lord, a função ogiva normal pela função logística, mais simples devido ao fato de não envolver integração, além de ser uma função explícita dos parâmetros do item e da habilidade. Rasch, independentemente, em 1960 propôs o modelo unidimensional de um parâmetro (dificuldade), usando a função ogiva normal e esse modelo foi mais tarde descrito por Wright em 1968, substituindo a ogiva normal pela função logística. Com a finalidade de obter mais

informação das respostas dos indivíduos, Samejima em 1969 propôs o modelo de resposta gradual. Bock em 1972, Andrich em 1978, Masters em 1982 e Muraki em 1992 também propuseram modelos para respostas politômicas, isto é, modelos para mais de 2 categorias de resposta, assumindo diferentes estruturas entre essas categorias. Bock e Zimowski em 1997 introduziram os modelos logísticos de um, dois e três parâmetros para duas ou mais populações de respondentes.

No Brasil, a primeira aplicação da TRI foi na análise das questões das provas do vestibular da UFRGS de 1993 (Albanese et al. 1993) e a partir desse pioneirismo, os órgãos governamentais vêm valorizando e difundindo o emprego desta técnica nas avaliações educacionais brasileiras, com o objetivo de produzir informações sobre o desempenho dos alunos, bem como sobre as condições internas e externas que interferem no processo de ensino-aprendizagem.

3.2.2 Modelos da Resposta ao Item

A TRI é um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da habilidade (traço latente) do respondente. Esse traço latente é uma variável que deve ser inferida a partir da observação de variáveis secundárias que estejam relacionadas a ela.

Os modelos de resposta ao item classificam-se de acordo com:

- a) Os itens: dicotômicos ou não;
- b) Número de parâmetros: 1, 2 ou 3;
- c) Número de habilidades ou traços latentes avaliadas: uma (modelos unidimensionais) ou mais (modelos multidimensionais).

Os modelos multidimensionais e politômicos podem ser encontrados em Linden e Hambleton (1997).

Os modelos apresentados nessa subseção podem ser utilizados tanto para a análise de itens de múltipla escolha dicotomizados (corrigidos como certo ou errado) quanto para a análise de itens abertos (de resposta livre), quando avaliados também de forma dicotomizada.

Os modelos logísticos para itens dicotômicos, empregados principalmente na educação, são os modelos de TRI mais utilizados, e diferenciam-se pelo número de

parâmetros utilizados para descrever o item: 1, 2 e 3 parâmetros, que consideram respectivamente, a dificuldade do item, a dificuldade e a discriminação, e a dificuldade, discriminação e probabilidade de acerto casual ou “chute”.

Será apresentado com maior detalhe o modelo logístico unidimensional de 3 parâmetros, visto que os outros dois modelos são obtidos facilmente a partir deste que é mais complexo.

3.2.2.1 Modelo Logístico de 3 Parâmetros

Classifica-se como um modelo que envolve apenas um traço latente. É utilizado quando o teste é aplicado a uma população de respondentes e os itens são dicotômicos ou dicotomizados, por exemplo, falso/ verdadeiro; certo/ errado.

Ao analisar I itens de n indivíduos, o modelo é dado por:

$$P(X_{ij} = 1/\theta_j) = c_i + \frac{1 - c_i}{1 + e^{-Da_i(\theta_j - b_i)}}; \quad i = 1, 2, \dots, I; j = 1, 2, \dots, n \quad (4)$$

onde:

$$X_{ij} = \begin{cases} 1, & \text{se o indivíduo } j \text{ responde corretamente ao item } i; \\ 0, & \text{caso contrário.} \end{cases}$$

θ_j : é a habilidade do j -ésimo indivíduo.

a_i : parâmetro de discriminação do item i .

b_i : parâmetro de dificuldade do item i .

c_i : parâmetro que representa a probabilidade de acerto casual do item i .

D : é uma constante (fator de escala) e é igual a 1. Quando se deseja que a função logística forneça resultados semelhantes ao da função ogiva normal utiliza-se o valor 1,7.

A equação (4) corresponde à probabilidade do indivíduo j responder corretamente ao item i dado que tem habilidade θ_j ; ou seja, a proporção de respostas corretas ao item i dentre todos os indivíduos com habilidade θ_j , no contexto educacional. Os parâmetros dos itens podem ser mais bem visualizados através de suas representações gráficas, conhecidas como curva característica do item (CCI), com inclinação e deslocamento na escala de habilidade definida pelos parâmetros. A forma de uma CCI descreve o quanto uma mudança no traço latente do indivíduo relaciona-se com a mudança na probabilidade de resposta específica.

Pela Figura 1 pode-se perceber que indivíduos com maior habilidade possuem maior probabilidade de acertar o item e que esta relação não é linear, isto é, possui uma representação em forma de “S” (sigmóide).

Os parâmetros do item se referem à:

a) Discriminação do item i : a_i

Este parâmetro é proporcional à derivada da tangente da curva no ponto de inflexão, ou seja, à inclinação da curva característica do item (CCI) no ponto b_i . Por este motivo, a_i também é chamado de parâmetro de inclinação do item i . É um parâmetro estritamente positivo, pois itens com a_i negativo indicaria que a probabilidade de responder corretamente ao item diminui com o aumento da habilidade. Valores muito altos indicam uma CCI muito íngreme, discriminando os indivíduos que possuem habilidade abaixo do parâmetro b_i e os que possuem habilidade acima do valor de b_i .

Uma questão com alto poder de discriminação indica que os candidatos com maior (ou menor) habilidade tenderão a acertar (ou errar) esta questão. Se a discriminação for baixa, aqueles com pouca ou muita habilidade terão probabilidades semelhantes de acertarem (ou errarem) a questão.

b) Dificuldade do item i : b_i

Este parâmetro, também chamado de parâmetro de posição ou locação do item i , é medido na mesma escala da habilidade e representa a habilidade necessária para uma probabilidade de acerto igual a $\frac{1+c_i}{2}$. Assim, quanto maior o valor de b_i , mais difícil é o item, e vice-versa.

c) Probabilidade de acerto casual: c_i

É o parâmetro que indica a probabilidade de acerto casual no item i , ou seja, a probabilidade de resposta correta dada por indivíduos de baixa habilidade (resposta aleatória). Como se trata de uma probabilidade, assume valores entre 0 e 1.

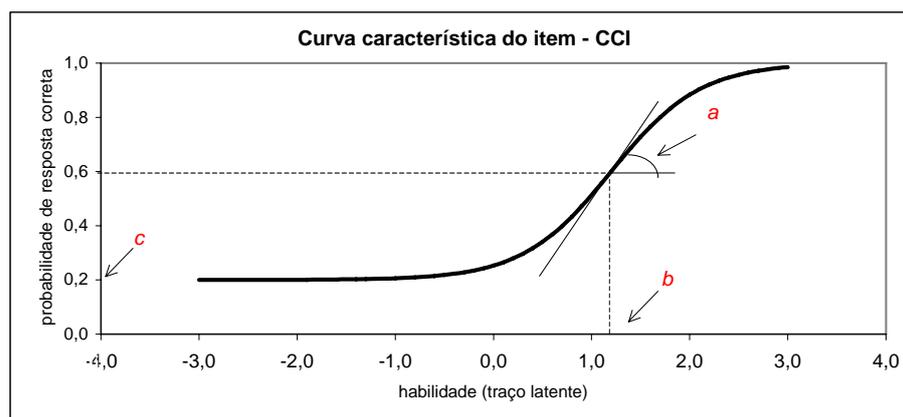


Figura 1. Curva Característica do Item para Modelo Logístico de 3 Parâmetros

As habilidades e os parâmetros dos itens são estimados a partir das respostas dos respondentes submetidos a estes itens, mas, uma vez estabelecida a escala de medida da habilidade, os valores dos parâmetros dos itens não mudam, ou seja, seus valores são invariantes a diferentes grupos de respondentes, desde que suas habilidades sejam medidas na mesma escala.

3.2.2.2 Modelos Logísticos de 1 e 2 parâmetros

Os dois outros modelos logísticos podem ser facilmente obtidos a partir do modelo logístico de 3 parâmetros. Por exemplo, quando não existe possibilidade de acerto ao acaso, isto é, $c = 0$, tem-se o modelo logístico unidimensional de 2 parâmetros:

$$P(X_{ij} = 1/\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}; \quad i = 1, 2, \dots, I; j = 1, 2, \dots, n \quad (5)$$

Se além de não existir resposta ao acaso e todos os itens tiverem o mesmo poder de discriminação, tem-se o modelo logístico de 1 parâmetro, também conhecido como Modelo de Rasch. Este modelo é dado por:

$$P(X_{ij} = 1 / \theta_j) = \frac{1}{1 + e^{-D(\theta_j - b_i)}}; \quad i = 1, 2, \dots, I; j = 1, 2, \dots, n \quad (6),$$

com os parâmetros já definidos anteriormente.

3.2.2.3 Outros Modelos de Teoria da Resposta ao Item

Existem outros modelos de Teoria da Resposta ao Item, os chamados modelos de respostas politômicas, nem os modelos de múltipla escolha, modelos de intervalos sucessivos e o modelo ordenado-parcionado de Wilson, além dos modelos multidimensionais e não-paramétricos, descritos por Embretson e Reise (2000), Lindsen e Hambleton (1997), Andrade et al. (2000) e Baker (2001).

Os modelos para respostas politômicas utilizam mais intensamente a informação contida nos questionários (testes), mas por conseqüência, necessitam de um número maior de parâmetros a serem estimados. Nessa categoria de modelos estão inclusos os modelos tanto para a análise de itens abertos (de resposta livre) quanto para a análise de itens de múltipla escolha que são avaliados de forma ordenada. O modelo de resposta gradual de Samejima é uma generalização do modelo logístico de 2 parâmetros, assumindo que as categorias de resposta de um item podem ser ordenadas entre si, como uma escala de Likert. Em 1990, Muraki desenvolveu uma modificação no modelo de resposta gradual, que facilitou o uso do mesmo na análise de questionários onde todos os itens têm o mesmo número de categorias de resposta, e estas devem ser igualmente espaçadas.

Outro modelo de TRI para respostas politômicas é o modelo de crédito parcial, desenvolvido por Masters em 1980. Trata-se de uma extensão do modelo de Rasch para mais de 2 categorias. E relaxando a hipótese de poder de discriminação constante para todos os itens, surgiu em 1992 o modelo de crédito parcial generalizado, proposto por Muraki. É um quinto modelo, que gera muita confusão

devido ao nome é o modelo de escala gradual, muito semelhante ao modelo de resposta gradual. A diferença está no fato de que cada item é descrito por um único parâmetro de locação de escala, que reflete a relativa facilidade ou dificuldade do particular item.

Uma síntese dos modelos de TRI encontra-se no Quadro 1.

Quadro 1. Modelos Unidimensionais de TRI

Natureza do Item	Dificuldade	Dificuldade e Discriminação	Dificuldade, discriminação e “chute”
Respostas Dicotômicas	Modelo Logístico de 1 parâmetro (Modelo de Rasch)	Modelo Logístico de 2 parâmetros	Modelo Logístico de 3 parâmetros
Respostas Politômicas	Modelo de Crédito Parcial Modelo de Escala Gradual Modelo de Resposta Gradual Modificada	Modelo de Resposta Gradual Modelo de Crédito Parcial Generalizado	

3.2.3 Função de Informação do Item

É uma medida bastante usada em conjunto com a CCI. Permite analisar quanto um item traz de informação para a medida de habilidade. A informação fornecida pelo item i no nível de habilidade θ_j é dada por:

$$I_i(\theta_j) = \frac{\left[\frac{\partial}{\partial \theta_j} P_i(\theta_j) \right]^2}{P_i(\theta_j) Q_i(\theta_j)} \quad (7)$$

com

$$P_i(\theta_j) = P(X_{ij} = 1/\theta_j) \quad (8)$$

$$Q_i(\theta_j) = 1 - P_i(\theta_j). \quad (9)$$

A cada item está associado um intervalo na escala da habilidade no qual o item tem maior poder de discriminação e as funções de informação dos itens são mais precisas. Este intervalo é definido em torno do valor do parâmetro b_i e corresponde à $(b_i - D; b_i + D)$.

3.2.4 Função de Informação do Questionário

A informação fornecida pelo questionário ou teste é dada pela soma das informações fornecidas para cada item do questionário:

$$I(\theta_j) = \sum_{i=1}^I I_i(\theta_j) \quad (10)$$

O erro padrão de medida (EP), chamado na TRI de erro padrão de estimação é dado por:

$$EP(\theta_j) = \frac{1}{\sqrt{I(\theta_j)}}. \quad (11)$$

O EP permite estabelecer intervalos de confiança em torno das habilidades dos sujeitos.

3.2.5 Escala de Habilidade

A escala da habilidade (traço latente) é uma escala arbitrária, onde o que importa são as relações de ordem entre seus pontos e não a magnitude. A habilidade pode teoricamente assumir qualquer valor real entre $-\infty$ e $+\infty$.

Para resolver esta indeterminação do modelo, é necessário estabelecer-se uma origem e uma unidade de medida para a definição da escala. A mais comum é a escala (0,1), onde o zero representa o valor médio das habilidades dos indivíduos da população em estudo e o 1, o desvio padrão das habilidades dos indivíduos da população em estudo. Neste caso, os valores dos parâmetros b_i variam entre -2 e 2. Em relação ao parâmetro a_i , esperam-se valores entre 0 e 2, sendo que os valores mais apropriados estariam entre 0,6 e 1,7, quando se utiliza $D = 1,7$.

3.2.6 Suposições do Modelo

Os modelos de TRI mais empregados fazem suposições relativamente simples, mas fortes sobre o relacionamento entre as respostas aos itens e o traço latente. As duas suposições básicas são a unidimensionalidade e a independência local. A unidimensionalidade é um dos pressupostos da maioria dos modelos de TRI. A unidimensionalidade foi, desde sempre, desejada num teste, mesmo na TCT, pois os conceitos de homogeneidade e dificuldade só fazem sentido quando se assume avaliar um único atributo (Lai et al. 2006).

A unidimensionalidade é a homogeneidade do conjunto de itens que supostamente devem estar medindo um único traço latente. Para satisfazer o postulado da unidimensionalidade é necessário admitir que haja uma habilidade (fator) dominante responsável pelo conjunto de itens, embora sabe-se que instrumentos de avaliação de qualidade de vida são multidimensionais por natureza e construção. Este fator dominante é o fator que se supõe estar sendo medido no questionário, isto é, o traço latente.

A dimensionalidade é verificada tipicamente por meio da análise fatorial, feita a partir da matriz de correlações tetracóricas, em especial a análise fatorial não linear. Outro procedimento sugerido é baseado no método de máxima

verossimilhança, muito embora a definição de unidimensionalidade de uma escala seja um assunto em discussão na literatura especializada.

A unidimensionalidade não pode ser completamente alcançada e alguns fatores podem atuar concomitantemente com o traço em avaliação. Assim, o desempenho num teste pode ser influenciado por outras variáveis cognitivas, pela personalidade do sujeito ou até por fatores inerentes à própria aplicação do teste, como a motivação, ansiedade, capacidade para trabalhar em velocidade ou uso correto das folhas de respostas, dentre outros. Portanto, a questão principal não é avaliar se um teste é estritamente unidimensional, mas sim em que medidas as dimensões adicionais geram distorções na medida principal do fator dominante a que os itens estão mais ou menos relacionados (Vitória et al. 2006). Na medida em que vários fatores podem afetar a unidimensionalidade de um teste, é também possível que um teste seja unidimensional para uma população, mas não para outra.

Outra situação particular e relevante que dificulta a unidimensionalidade do teste refere-se ao formato dos respectivos itens, que devem sempre estar na mesma direção (itens estritamente negativos ou positivos). Muitas vezes a inversão dos itens pode produzir duas dimensões relativamente independentes ao invés de dois pólos de uma mesma dimensão. Os instrumentos elaborados pela OMS, WHOQOL-100 e WHOQOL-BREVE, são exemplos de instrumentos com itens em duas direções.

A independência local ou condicional postula que para uma dada habilidade, as respostas aos diferentes itens do instrumento são independentes, o que é fundamental para estimar os parâmetros do modelo. A suposição de independência local é análoga à suposição na TCM de erros não correlacionados condicionais aos escores verdadeiros do examinado.

Ainda que pareça improvável que os componentes de um mesmo instrumento não sejam correlacionados, a independência local assegura que, se existir correlação, esta se deve à influência de outros fatores que não o dominante. Se estes outros fatores forem mantidos constantes, o fator dominante será a exclusiva fonte de variação, e as respostas se tornam independentes. A violação de independência local, chamada dependência local, ocorre quando as respostas aos itens não dependem apenas dos seus traços latentes, mas das suas respostas aos demais itens dos

questionários, o que é muito provável. Este fenômeno pode ocorrer em muitas situações de testes educacionais e de personalidade.

A presença de dependência local pode trazer sérias conseqüências em relação à aplicabilidade dos modelos unidimensionais de TRI, exercendo papel fundamental na computação da função de verossimilhança.

A unidimensionalidade implica independência local, então se a suposição da unidimensionalidade for satisfeita, esta também será. Assim, itens devem ser elaborados de modo a satisfazer a suposição de unidimensionalidade.

3.2.7 Métodos de Estimação

Um ponto crítico na TRI é a estimação dos parâmetros envolvidos nos modelos, em particular quando necessita-se estimar tanto os parâmetros dos itens quanto as habilidades. Basicamente, têm-se a estimação das habilidades, quando já se conhecem os parâmetros dos itens, estimação dos parâmetros dos itens quando já se conhecem as habilidades e a estimação conjunta das habilidades e dos parâmetros dos itens. Na TRI o processo de estimação dos parâmetros dos itens é conhecido como calibração. Inicialmente, a estimação era feita através do método da máxima verossimilhança conjunta que envolve um número muito grande de parâmetros a serem estimados simultaneamente e, conseqüentemente, grandes problemas computacionais. Em 1970, Bock e Lieberman introduziram o método da máxima verossimilhança marginal para a estimação dos parâmetros em duas etapas. Na primeira etapa estimam-se os parâmetros dos itens, assumindo-se uma certa distribuição para as habilidades. Na segunda etapa, assumindo os parâmetros dos itens conhecidos, estimam-se as habilidades. Apesar do avanço que esse método trouxe para o problema, ele requeria que os parâmetros fossem estimados simultaneamente. Em 1981, Bock e Aitkin propuseram uma modificação no método acima, utilizando o algoritmo EM (Esperança e Maximização) de modo a permitir que os itens pudessem ter seus parâmetros estimados em separado, facilitando em muito o aspecto computacional do processo de estimação. Mais recentemente, métodos bayesianos foram propostos para, entre outras coisas, resolver o problema de estimação dos parâmetros dos itens (Presa 2005; Valle 1999; Cúri 2006; Vergara 2005).

De acordo com Valle (1999), vários autores sugerem que cada respondente tenha sido submetido a pelo menos 30 itens e que cada item tenha sido submetido a pelo menos 300 respondentes, visando melhores estimativas com pequenos erros padrões.

3.2.7.1 Estimação das Habilidades dos Indivíduos

Neste caso, os parâmetros dos itens são conhecidos. Utiliza-se quando se deseja submeter indivíduos a itens já calibrados, visando à estimação de suas habilidades com objetivos de classificação ou seleção. É bastante utilizado, o que faz com que aumente o interesse na criação de bancos de itens, uma das grandes contribuições da TRI.

As habilidades dos indivíduos submetidos ao teste podem ser estimadas a partir do método da máxima verossimilhança ou métodos bayesianos, através do estimador bayesiano esperança a posteriori (EAP) ou pelo estimador modal de Bayes (MAP). Ambos métodos exigem procedimentos iterativos como o método de Newton-Raphson que envolvem cálculos complexos e, conseqüentemente, recursos computacionais especializados.

O método da máxima verossimilhança (MV) é um dos métodos mais populares de estimação, devido às suas boas propriedades, como: consistência das estimativas, isto é, convergência da estimativa ao valor verdadeiro em grandes amostras; eficiência, ou seja, erros padrões relativamente menores e com distribuição normal. Além disso, a verossimilhança completa dos dados tem uma relação com a distribuição quiquadrado, o que permite a comparação dos modelos pelo teste da razão de verossimilhança. Embora tenha essas propriedades desejáveis, o método da MV não é sustentado para alguns modelos de TRI, nem ao aumentar-se o tamanho amostral.

O problema da não existência de estimadores de MV em algumas situações pode ser resolvido por procedimentos bayesianos, que consistem em incorporar qualquer informação a priori de modo a modificar a função de verossimilhança. O estimador bayesiano EAP (esperança a posteriori) é a média da distribuição a posteriori de θ (traço latente) dado as respostas aos itens. Ele está definido para

qualquer padrão de resposta e tem um erro médio menor do que qualquer outro estimador. O estimador modal de Bayes (MAP) consiste no valor de θ que maximiza o logaritmo natural da distribuição a posteriori, que é o produto da função de verossimilhança e da priori, e apresenta características semelhantes ao EAP.

3.2.7.2 Estimação dos Parâmetros dos Itens

Neste caso, as habilidades dos respondentes são conhecidas, o que na prática não ocorre. Os parâmetros dos itens são estimados item a item, devido à complexidade da estimação de todos os itens conjuntamente. Para obter os parâmetros dos I itens do teste, repete-se o processo independentemente I vezes.

Os parâmetros dos itens podem ser estimados pela máxima verossimilhança marginal. O método da máxima verossimilhança marginal, proposto por Bock e Lieberman em 1970, apresenta algumas vantagens em relação aos outros métodos. A idéia básica deste procedimento é a integração em θ , de modo que a função de verossimilhança não dependa dos parâmetros de habilidade. Também é necessário utilizar um método iterativo, geralmente o método de Newton-Raphson na sua forma multivariada.

3.2.7.3 Estimação conjunta dos Parâmetros dos Itens e das Habilidades dos Indivíduos

É a situação mais comum: têm-se I itens submetidos a n indivíduos e desejam-se estimar tanto os parâmetros dos itens quanto suas habilidades.

Estimam-se ao mesmo tempo os parâmetros dos itens e as habilidades e devido à dificuldade de estimarem-se conjuntamente os parâmetros dos itens e as habilidades, procede-se em duas fases.

Na primeira fase, os parâmetros dos itens são estimados a partir do método da máxima verossimilhança conjunta, do método da máxima verossimilhança marginal ou de métodos bayesianos.

Uma vez estimados os parâmetros dos itens, na fase 2, os parâmetros de habilidade dos respondentes são estimados considerando-se os parâmetros dos itens

como conhecidos. Assim, estimam-se as habilidades na mesma escala dos parâmetros dos itens através dos métodos da máxima verossimilhança ou bayesianos.

O método da máxima verossimilhança conjunta (MVC) foi o primeiro método a ser utilizado nesse tipo de problema de estimação dos parâmetros dos itens, talvez por ser o mais natural. É um método bastante complexo, pois envolve a resolução simultânea das estimativas para todos os parâmetros. Aqui também pode ocorrer o processo de indeterminação e à medida que o tamanho amostral aumenta, mais parâmetros necessitam ser estimados, o que faz com que as propriedades assintóticas dos estimadores de máxima verossimilhança não sejam válidas nesse caso.

O método da máxima verossimilhança marginal (MVM), ao contrário do MVC, que não faz qualquer suposição sobre a distribuição da habilidade, assume que os respondentes representam uma amostra aleatória de uma população na qual a habilidade é distribuída segundo uma determinada função densidade. A essência deste procedimento é a integração em θ , de modo que a função de verossimilhança não dependa dos parâmetros de habilidade. E conseqüentemente, os parâmetros dos itens são estimados na distribuição marginal, e essa estimação não depende mais da estimação das habilidades dos respondentes, mas sim das distribuições dessas habilidades. Para esse método não há o problema de indeterminação, pois estabelece uma distribuição a priori, que é geralmente uma distribuição normal com média 0 e desvio padrão 1.

O MVM é trabalhoso e inapropriado quando se tem muitos itens. Este problema é solucionado com o algoritmo EM, que é um processo iterativo para determinação de estimativas de máxima verossimilhança de parâmetros de modelos de probabilidade na presença de variáveis aleatórias não-observadas. Esse algoritmo produz estimativas consistentes para os parâmetros dos itens e é computacionalmente mais simples. O algoritmo é dividido em duas etapas: esperança (E) e maximização (M). Em cada ciclo do algoritmo, estimativas dos parâmetros dos itens são calculadas em uma escala definida a partir da normalização e reescalonamento da distribuição a posteriori da habilidade, que faz com que parâmetros de locação e escala da distribuição a posteriori tenham mesmos valores dos correspondentes da distribuição a priori. No fim, são fornecidas estimativas de máxima verossimilhança dos

parâmetros dos itens e a estimativa da distribuição a posteriori das habilidades, na mesma escala. Algumas vezes esse procedimento não fornece as melhores estimativas, geralmente na estimação do ML-3 (parâmetro do acerto casual) e na estimação da dificuldade com valores extremos (itens com dificuldades muito grandes ou muito pequenas).

Portanto, para o processo de estimação ser bem sucedido, é importante ter-se respondentes com habilidades cobrindo todo o espectro do conhecimento a ser avaliado. Em situações problemáticas devem ser aplicados procedimentos bayesianos a partir da incorporação de prioris para os parâmetros dos itens, pois eles fornecem estimativas para todos os itens e habilidades.

Existem outros métodos de estimação que não foram abordados. Mais informações sobre os métodos de estimação consultar Andrade et al. (2000).

3.2.8 Ajuste do Modelo

O ajuste do modelo constitui outro ponto fraco da TRI. Há uma série de testes e estatísticas de ajuste, mas nenhuma se constitui no padrão-ouro ou referencial. Há um consenso entre os pesquisadores da área de que itens com ajustes individuais ruins, verificado geralmente através dos testes qui-quadrados implementados nos softwares disponíveis, não comprometem o instrumento como um todo. Os modelos da resposta ao item só podem ser considerados vantajosos quando o ajuste do modelo aos dados de interesse for satisfatório. Um modelo mal-ajustado não fornecerá parâmetros invariantes para os itens e para as habilidades.

Há basicamente dois procedimentos gerais para estimar o ajuste do item, que são os procedimentos gráficos e os testes formais. O primeiro consiste em comparar graficamente as curvas de resposta ao item estimada e observada, onde ao longo do traço latente se podem perceber as discrepâncias ou resíduos. Ajustes pobres podem ser devidos à multidimensionalidade inerente do instrumento, resultando na escolha equivocada de modelos unidimensionais, escolha do melhor modelo dicotômico ou politômico, não monotonicidade das relações item-traço latente e o fato de um grupo heterogêneo de examinados ter sido retirado de outra população e comprometer o ajuste, além de uma fraca construção do item.

Os testes formais encontram-se em Embretson e Reise (2000). Um destes testes é o Qui-quadrado. Para o cálculo dos valores de qui-quadrado associados a cada um dos modelos, os sujeitos são agrupados em subgrupos segundo habilidades semelhantes e progressivamente mais altas. Esses valores são obtidos calculando-se a probabilidade teórica para cada subgrupo a partir da CCI e os resíduos relativos à probabilidade real observada.

3.3 Avaliação de Qualidade de Vida

Até há pouco tempo, qualidade de vida referia-se às condições biológicas e sócio-econômicas mínimas, capazes de atender às necessidades de sobrevivência da própria espécie humana. A preocupação do ser humano era suprir suas necessidades básicas, como moradia, alimentação, segurança, transporte, saúde, etc..., sem se preocupar com outros aspectos que hoje são considerados tão importantes: lazer, atividade física, convívio social, satisfação pessoal, auto-realização, dentre outros (Cruz 2005; Minayo et al. 2000).

Avaliações de qualidade de vida têm sido cada vez mais incorporadas às práticas médicas. Nos últimos 50 anos, vem crescendo a preocupação dos profissionais da área da saúde em avaliar a efetividade dos tratamentos médicos e o impacto das diversas doenças através da perspectiva do próprio paciente, o que representa uma grande mudança no enfoque apenas da ausência de doença, tão comum antigamente. As investigações deixaram de ser baseadas apenas nas impressões dos médicos e nos parâmetros fisiológicos para então levar em conta a impressão do paciente acerca de sua própria saúde. Tradicionalmente, o atendimento médico era focalizado no diagnóstico e tratamento, e os resultados eram medidos através de indicadores de morbidade e mortalidade.

A importância da avaliação de qualidade de vida aumentou bruscamente tornando-se um atributo importante da investigação clínica e na formulação de políticas de saúde pública. A qualidade de vida percebida pelo paciente permite mensurar o verdadeiro impacto de uma enfermidade e assim, avaliar os diversos domínios da mesma, como por exemplo, aspectos físicos, emocionais, desempenho social, funcionamento no dia-a-dia, dentre outros.

De acordo com Cruz (2005), a medida de QV é imperativa, uma vez que insere os indivíduos nas avaliações relacionadas à sua própria saúde e tratamentos. E essa medida vem sendo cada vez mais utilizada como desfecho em ensaios clínicos e como medida de efetividade e de qualidade dos cuidados em saúde. Há evidências de que a QV seja uma medida válida e confiável, mesmo que baseada em instrumentos reduzidos e de mais fácil administração e compreensão.

O grupo de especialistas em qualidade de vida da OMS, em 1995, definiu qualidade de vida como “a percepção do indivíduo de sua posição na vida, no contexto do sistema cultural e de valores em que ele vive e em relação a seus objetivos, expectativas, padrões e preocupações”. E “embora não haja definição consensual de qualidade de vida, há concordância considerável entre os pesquisadores acerca de algumas características do construto qualidade de vida”. São citadas três características: subjetividade, multidimensionalidade e bipolaridade. Quanto à subjetividade, não há dúvidas, pois há de se diferenciar as condições objetivas (por exemplo, recursos materiais) das subjetivas (por exemplo, satisfação com os recursos). O WHOQOL GROUP também aponta para a mesma direção, ao propor que as questões do instrumento de avaliação de qualidade de vida se preocupem com as percepções das pessoas, em diferentes níveis de questionamento. Segundo esse grupo de experts da OMS, devem-se fazer questões acerca de avaliações globais dos comportamentos, estados e capacidades, pois esses questionamentos é que trazem as informações sobre a qualidade de vida, enquanto o relato do funcionamento traz informações sobre o estado de saúde.

Não importando como seja avaliada, a qualidade de vida é de suma importância para garantir um conhecimento mais amplo do efeito da doença no paciente e assim, melhorar os tratamentos clínicos já existentes, especialmente, para as doenças crônicas.

3.3.1 Instrumentos de avaliação de qualidade de vida

A qualidade de vida é geralmente medida através de uma coleção complexa de itens, escalas, domínios e instrumentos. (Cruz 2005; Lohr 2002)

Um item é uma questão única, por exemplo, “Como você avaliaria sua qualidade de vida?”.

Uma escala contém as categorias ou outros mecanismos utilizados para expressar a resposta à questão. Por exemplo, as opções para a questão citada acima poderiam ser: “Muito ruim”, “Ruim”, “Nem ruim nem boa”, “Boa”, “Muito boa”.

Um domínio ou dimensão é um bloco de itens relacionados, que se referem a uma determinada área de comportamento ou experiência que está sendo medida. Por exemplo, saúde mental ou capacidade funcional. Um domínio também pode ser representado por um único item.

Um instrumento ou índice é o conjunto de itens usados para obter os dados desejados. Um instrumento pode conter uma única questão global ou múltiplos itens que podem ou não ser categorizados em domínios específicos.

Os resultados gerados pelo instrumento podem ser apresentados de duas formas, uma em que o escore é computado por domínio e são citados, individualmente, um após o outro, para formar um perfil, ou os domínios são agregados para formar um escore global único.

Os instrumentos também podem ser categorizados de acordo com a perspectiva que eles se propõem a avaliar, seja elas de qualidade de vida geral, qualidade de vida relacionada à saúde ou relacionada a alguma doença em específico. Os instrumentos utilizados nesse trabalho, apresentados com detalhe a seguir, são instrumentos gerais de avaliação de qualidade de vida.

É indiscutível a exigência de se obedecer determinados critérios científicos no desenvolvimento dos instrumentos de avaliação, o que passa pela necessidade de possuírem certas propriedades. Em primeiro lugar, um bom instrumento deve ser discriminativo, isto é, diferenciar as pessoas com melhor QV das com pior QV. E se o foco for detectar mudanças na qualidade de vida através do tempo ou como efeito de algum tratamento, diz-se que o instrumento é avaliativo. Recomenda-se também que o instrumento seja preditivo, isto é, se o objetivo é utilizar uma medida alternativa às tradicionais medidas fisiológicas para avaliação de desfechos, a medida de qualidade de vida pode ser uma variável preditora de desfechos importantes.

A escolha de um instrumento depende do propósito de estudo. Pode-se usar um único instrumento ou uma bateria deles em conjunto.

3.3.1.1 Instrumentos de Avaliação de Qualidade de Vida desenvolvidos pela Organização Mundial de Saúde

A ausência de um instrumento que avaliasse a qualidade de vida *per se*, com uma perspectiva genuinamente internacional, fez com que a OMS constituísse um grupo de Qualidade de Vida (Grupo WHOQOL) com a finalidade de desenvolver instrumentos capazes de fazê-lo dentro de uma perspectiva transcultural, devido ao fato de envolver vários centros colaborativos.

O passo inicial foi a clarificação do conceito de qualidade de vida. Os demais passos e etapas estão sintetizados no Quadro 2. As etapas são descritas com detalhes nos documentos da OMS.

Quadro 2. Estágios do desenvolvimento do WHOQOL-100

Estágio	Método	Produto	Objetivos
1) Clarificação do conceito	Revisão por experts internacionais	- Definição de qualidade de vida - Definição de um protocolo para o estudo	Estabelecimento de um consenso para uma definição de qualidade de vida e para uma abordagem internacional da avaliação de qualidade de vida
2) Estudo piloto qualitativo	- Revisão por experts - Grupos focais - Painel escrito de experts e leigos	- Definição de domínio e subdomínios - Elaboração de um conjunto de questões	Exploração do conceito de qualidade de vida através das culturas e geração de questões
3) Desenvolvimento de um Piloto	Administração do WHOQOL piloto em 15 centros para 250 pacientes e 50 normais	- Padronização de um questionário de 300 questões	Refinamento da estrutura do WHOQOL. Redução do conjunto de questões.
4) Teste de campo	Aplicação em grupo homogêneos de pacientes	- Estrutura comum de domínios - Conjunto de 100 questões - Escala de respostas equivalentes nas diferentes línguas	Estabelecimento de propriedades psicométricas do WHOQOL.

Trata-se de um instrumento que reflete a multidimensionalidade própria do termo qualidade de vida. É composto por 96 itens divididos em 6 domínios: físico, psicológico, nível de independência, relações sociais, meio-ambiente e espiritualidade/religião/crenças pessoais (Fleck et al. 2003), além de 4 questões sobre qualidade de vida em geral. Cada domínio é constituído por facetas que são avaliadas por quatro questões, distribuídos em uma escala com 5 categorias. Assim, o instrumento é composto por 24 facetas específicas e uma faceta geral que inclui questões de avaliação global de qualidade de vida. As questões do WHOQOL-100 foram formuladas para uma escala de respostas do tipo Likert, com 5 alternativas para cada escala de intensidade (nada - extremamente), capacidade (nada – completamente), frequência (nunca – sempre) e avaliação (muito insatisfeito – muito satisfeito; muito ruim – muito bom). Foi escolhida uma metodologia cuidadosa para selecionar as palavras que compõem as escalas em cada idioma, com a finalidade de manter a equivalência nas diferentes línguas. O Quadro 3 apresenta os domínios e facetas do WHOQOL-100.

Quadro 3. Domínios e facetas do WHOQOL-100

<p><i>Domínio I – domínio físico</i></p> <ol style="list-style-type: none"> 1. Dor e desconforto 2. Energia e fadiga 3. Sono e repouso <p><i>Domínio II - domínio psicológico</i></p> <ol style="list-style-type: none"> 4. Sentimentos 5. Pensar, aprender, memória e concentração 6. Auto-estima 7. Imagem corporal e aparência 8. Sentimentos negativos <p><i>Domínio III – nível de independência</i></p> <ol style="list-style-type: none"> 9. Mobilidade 10. Atividades da vida cotidiana 11. Dependência de medicação ou de tratamentos 12. Capacidade de trabalho <p><i>Domínio IV – Relações sociais</i></p> <ol style="list-style-type: none"> 13. Relações pessoais 14. Apoio social 15. Atividade sexual <p><i>Domínio V – ambiente</i></p> <ol style="list-style-type: none"> 16. Segurança física e proteção 17. Ambiente no lar 18. Recursos financeiros 19. Cuidados de saúde e sociais: disponibilidade e qualidade 20. Oportunidades de adquirir novas informações e habilidades 21. Participação em, e oportunidades de recreação/habilidades 22. Ambiente físico: (poluição/ruído/trânsito/clima) 23. Transporte <p><i>Domínio VI – aspectos espirituais/religião/crenças pessoais</i></p> <ol style="list-style-type: none"> 24. Espiritualidade/religião/crenças pessoais

As características psicométricas do WHOQOL-100 foram estabelecidas a partir de uma amostra de 8924 indivíduos provenientes de 19 centros colaborativos. Foi avaliada a consistência interna através do alfa de Cronbach. A validade discriminante para os itens foi avaliada através do teste *t* para comparar controles e doentes. A confiabilidade teste-reteste foi avaliada pela correlação de Pearson e a contribuição dos escores dos domínios por regressão múltipla. O WHOQOL-100 no seu estudo piloto apresentou boa consistência interna, quer se tomem as 100 questões ou as 24 facetas, ou ainda os 6 domínios, além de boa validade discriminante para doentes e normais. A escala de Beck para Depressão (BDI) e a escala de desesperança de Beck (BHS) foram utilizados como instrumentos de avaliação indireta de validade concorrente. A versão em português do WHOQOL-100 foi

desenvolvida no Departamento de Psiquiatria e Medicina Legal da Universidade Federal do Rio Grande do Sul pelo grupo do Dr. Marcelo Fleck (Fleck et al. 1999; Fleck et al. 2000).

A necessidade de instrumentos curtos que demandem pouco tempo para seu preenchimento, mas com características psicométricas satisfatórias, fez com que o Grupo de Qualidade de Vida da OMS desenvolvesse uma versão abreviada do WHOQOL-100, o WHOQOL-bref (WHOQOL GROUP, 1998). A versão abreviada também foi coordenada por Fleck et al. no Brasil. A aplicação do teste de campo foi em uma amostra de 300 indivíduos, sendo 50 voluntários-controles e 250 pacientes do Hospital de Clínicas de Porto Alegre, 125 internados e 125 ambulatoriais (Fleck et al. 2000).

O WHOQOL-BREVE consta de 26 questões, sendo duas questões gerais de qualidade de vida e as demais 24 representam cada uma das 24 facetas que compõem o instrumento original WHOQOL-100. Assim, diferente do WHOQOL-100 em que cada uma das 24 facetas é avaliada a partir de 4 questões, no WHOQOL BREVE é avaliada por apenas uma questão. Os dados que deram origem à versão abreviada foram extraídos do teste de campo de 20 centros em 18 países diferentes.

O critério de seleção das questões foi tanto psicométrico como conceitual. No nível conceitual, foi definido pelo Grupo de Qualidade de Vida da OMS de que o caráter abrangente do instrumento deveria ser preservado. Assim, cada uma das 24 facetas que compõem o instrumento original (WHOQOL-100) deveria ser representada por uma questão. No nível psicométrico foi então selecionada a questão mais altamente correlacionada com o escore total, calculado pela média de todas as facetas. Após esta etapa, os itens selecionados foram examinados por um painel de experts para estabelecer se representavam conceitualmente cada domínio de onde as facetas provinham. Dos 24 itens selecionados, seis foram substituídos por questões que definissem melhor a faceta correspondente.

Uma análise fatorial confirmatória foi realizada para uma solução a quatro domínios. Assim o WHOQOL-BREVE é composto por 4 domínios: Físico, Psicológico, Relações Sociais e Meio- ambiente. O Quadro 4 apresenta os domínios e facetas do instrumento abreviado.

Quadro 4. Domínios e facetas do WHOQOL – BREVE

<p><i>Domínio 1 - domínio físico</i></p> <ol style="list-style-type: none"> 1. Dor e desconforto 2. Energia e fadiga 3. Sono e repouso 9. Mobilidade 10. Atividades da vida cotidiana 11. Dependência de medicação ou de tratamentos 12. Capacidade de trabalho <p><i>Domínio 2 – domínio psicológico</i></p> <ol style="list-style-type: none"> 4. Sentimentos positivos 5. Pensar, aprender, memória e concentração 6. Auto-estima 7. Imagem corporal e aparência 8. Sentimentos negativos 24. Espiritualidade/religião/crenças pessoais <p><i>Domínio 3 – relações sociais</i></p> <ol style="list-style-type: none"> 13. Relações pessoais 14. Suporte (Apoio) social 15. Atividade sexual <p><i>Domínio 4 – meio ambiente</i></p> <ol style="list-style-type: none"> 16. Segurança física e proteção 17. Ambiente no lar 18. Recursos financeiros 19. Cuidados de saúde e sociais: disponibilidade e qualidade 20. Oportunidade de adquirir novas informações e habilidades 21. Participação em, e oportunidades de recreação/lazer 22. Ambiente físico: (poluição/ruído/trânsito/ clima) 23. Transporte

A versão em português também apresentou boa consistência interna, validade discriminante, validade concorrente, validade de conteúdo e confiabilidade teste-reteste. O WHOQOL-BREVE, assim como o WHOQOL-100, é um instrumento genérico, auto-aplicável, mas também pode ser administrado pelo entrevistador.

3.3.1.2 Escala de Qualidade de Vida de Flanagan

A Escala de Qualidade de Vida de Flanagan (EQVF) foi desenvolvida no âmbito dos Estados Unidos por Flanagan em 1982, e trata-se de uma escala psicométrica de natureza ordinal em cuja aplicação obtêm-se respostas para quinze itens agrupados em cinco dimensões: bem estar físico e material; relações com outras pessoas; atividades sociais, comunitárias e cívicas; desenvolvimento pessoal e

realização e recreação. Flanagan desenvolveu esta escala com os dados colhidos de uma amostra de três mil homens e mulheres norte-americanos, com idades de 30, 50 e 70 anos. Foi considerado um instrumento multidimensional e com evidência de validade de construto em função dos resultados da análise fatorial feita. É um instrumento breve de avaliação da qualidade de vida em geral. Burckhardt, em 1989, mudou essa pontuação para 7 respostas, com o objetivo de aumentar o espectro das mesmas. Essa escala foi inicialmente utilizada no Brasil por Hashimoto et al. (1996), citado por Santos et al. (2002), que a traduziram e aplicaram-na em pacientes ostomizados. Não há referência quanto à sua confiabilidade após a tradução. Apesar dessas limitações, a EQVF tem sido amplamente utilizada por ser de fácil aplicação e interpretação (Santos et al. 2002; Lentz et al. 2000).

As dimensões são mensuradas através de quinze itens onde o respondente tem sete opções de resposta, que vai de "muito insatisfeito" (escore 1) até "muito satisfeito" (escore 7). A pontuação máxima alcançada na avaliação da qualidade de vida proposta por Flanagan é de 105 pontos e a mínima de 15 pontos, que refletem baixa qualidade de vida. Cabe destacar que a escala é auto-aplicável. O Quadro 5 apresenta os itens e dimensões da EQVF.

Quadro 5. Dimensões/Itens da Escala de Qualidade de Vida de Flanagan

<p><i>1. Bem estar físico e material</i> Conforto material Saúde</p> <p><i>2. Relações com outras pessoas</i> Relação com parentes Ter e criar filhos Íntimo Amigos</p> <p><i>3. Atividades sociais, comunitárias e cívicas</i> Ajuda Participação</p> <p><i>4. Desenvolvimento pessoal e realização</i> Aprendizagem Auto-conhecimento Trabalho Comunicação</p> <p><i>5. Recreação</i> Socialização “fazer amigos” Recreação passiva Recreação ativa</p>
--

3.4 Avaliação de Qualidade de Vida em Idosos

O envelhecimento populacional é um dos maiores desafios da saúde pública contemporânea. Inicialmente ocorreu em países desenvolvidos, mas é nos países em desenvolvimento que o envelhecimento tem ocorrido de forma mais acentuada. No Brasil, o número de idosos com idades iguais ou superiores a 60 anos passou de 3 milhões em 1960 para 7 milhões em 1975, 14 milhões em 2002 e estima-se que alcançará 18 milhões em 2010. Considerando a continuidade das tendências verificadas para as taxas de fecundidade e longevidade da população brasileira, as estimativas para os próximos 20 anos indicam que a população idosa poderá exceder 30 milhões ao final deste período, chegando a representar quase 13% da população (Moraes 2004).

Um dos maiores feitos da humanidade foi a ampliação do tempo de vida, que se fez acompanhar de uma melhora substancial dos parâmetros de saúde das populações, ainda que estas conquistas estejam longe de se distribuírem de forma igualitária em diferentes países e contextos sócio-econômicos. E junto com essa conquista, vem o maior desafio: envelhecer com qualidade de vida, isto é, adicionar qualidade de vida aos anos que restam a esses indivíduos (Kalache et al. 1987; Pereira et al. 2005; Cassiano et al. 2005; Fleck et al. 2003; Fleck et al. 2006; Farquhar 1995; Ramos 2003; Chaimovicz 1997; McHorney 2003; Ware 2003).

No Brasil, a situação é precária: o envelhecimento caracteriza-se pelo acúmulo de incapacidades progressivas nas suas atividades funcionais e de vida diária, associada a condições sócio-econômicas adversas. E esse acelerado ritmo de envelhecimento brasileiro transforma-se num desafio em especial para a sociedade brasileira contemporânea, onde esse processo ocorre num cenário de profundas transformações econômicas, sociais, urbanas, industriais e familiares.

A tendência ao envelhecimento populacional está acarretando mudanças profundas em todos os setores da sociedade. A transição demográfica tem um crescente e profundo impacto em todos os âmbitos da sociedade, mas é na saúde que tem maior transcendência, tanto por sua repercussão nos diversos níveis assistenciais como pela demanda por novos recursos e estruturas (Ramos et al. 1987; Veras et al. 1987).

Assim, destaca-se a relevância científica e social de se investigar as condições que interferem no bem-estar e nos fatores associados à qualidade de vida dos idosos, no intuito de criar alternativas de intervenção e propor políticas na área da saúde, visando atender às demandas desse seguimento que cresce cada vez mais (Xavier et al. 2003, Berlim e Fleck 2003, Vecchia et al. 2005, Sousa et al. 2003; Veras 2003).

A avaliação do estado de saúde (Paixão e Reichenheim 2005) está diretamente ligada à qualidade de vida, influenciada por sexo, escolaridade, idade, situação econômica e presença de incapacidades (Iezzoni 2006). Quando se investiga a qualidade de vida relacionada à saúde em sua multidimensionalidade, identificam-se os principais aspectos a serem considerados em relação às potencialidades e peculiaridades que a vida do idoso tem, interferindo no seu processo saúde-doença.

Portanto, se o estudo e a avaliação da qualidade de vida tem mostrado sua importância em vários segmentos e disciplinas, elas também se revelam importantes para a população idosa. As sociedades ocidentais incorporaram vários preconceitos, que dificultam a vida do idoso e muitos desses foram incorporados pelos próprios idosos que se subvalorizam, como o fato de seus valores serem antiquados e assim inferiores, serem menos habilidosos para julgarem e criticarem, terem menos valor que os jovens, necessitando então de menos atenção, e a redução exponencial da capacidade de aproveitar a vida. Avaliar qualidade de vida em idosos é diferente do que para as demais faixas etárias. Na Gerontologia Social e na Pesquisa de Indicadores Sociais, uma definição comumente usada é de que qualidade de vida descreve quatro dimensões subjetivas ao seu conceito, duas das quais objetivas e duas que refletem o julgamento pessoal do indivíduo; saúde geral e estado funcional; estado sócio-econômico; satisfação de vida e auto-estima. Não são dimensões completas, mas são quatro dimensões centrais de um número infinito de aspectos de qualidade de vida (Farquhar 1995).

Em pesquisas de serviço de saúde, saúde geral (ambos bem-estar físico e psicológico) e estado funcional são julgados como importantes dimensões de qualidade de vida, especialmente para idosos, juntamente com suas altas taxas de doenças crônicas. Há poucos instrumentos usados e validados para avaliar qualidade de vida em idosos, especialmente no Brasil. Um deles é o WHOQOL-OLD, um

instrumento complementar para a medição de qualidade de vida em idosos, que pode fornecer informações adicionais a respeito da QV. É composto por 24 itens divididos em 6 facetas, relativas ao funcionamento dos sentidos, autonomia, atividades passadas, presentes e futuras, participação social, morte e morrer, e intimidade (Fleck et al. 2006, Fleck et al. 2003).

Aferir qualidade de vida é muito complexo e os domínios que requerem medida entre os idosos incluem os problemas de saúde, que podem levar à incapacidade e invalidez, saúde mental, habilidade funcional, estado geral de saúde, satisfação de vida, estado de espírito, controle (autonomia) e suporte social. Isso mostra a enorme variedade de dimensões a serem aferidas numa avaliação de qualidade de vida em idosos, levando à necessidade de instrumentos multidimensionais que possam captar a enorme variabilidade dos diferentes grupos de idosos. Tais instrumentos devem levar em conta as especificidades dessa faixa, como seus valores e experiências de vida, fatores ligados à saúde e diversas situações como aposentadoria, viuvez, dependência, perda de autonomia e de papéis sociais. Todos esses fatores aumentam a complexidade de tal avaliação. Quando se avalia qualidade de vida em idosos, geralmente usa-se um instrumento de uso universal, construído e validado para diversas faixas de idade e populações (Paschoal 2000).

Segundo Paschoal (2000), não foi encontrado nenhum instrumento genuinamente brasileiro para avaliar qualidade de vida em idosos. Doll em 1998, citado por Paschoal, utilizou o ISS (Índice de Saúde Subjetiva) associado a outro instrumento americano para avaliar satisfação de vida de pessoas com 50 anos ou mais, comparando idosos brasileiros com idosos alemães. O ISS abrange 6 dimensões de saúde subjetiva de pessoas idosas, como doenças crônicas, doenças agudas, dor, lista de doenças e problemas, auto-avaliação e aspecto funcional. Em 1999, Gonçalves et al. avaliaram a qualidade de vida de idosos em uma amostra de idosos frequentadores de uma universidade aberta da terceira idade em João Pessoa (Paraíba) usando a Escala de Qualidade de Vida de Flanagan (EQVF) (Santos et al. 2002). E em 2005, Fleck e colaboradores aplicaram o WHOQOL-OLD (versão em português) numa amostra de 424 idosos de Porto Alegre, e concluíram que o WHOQOL-OLD é um instrumento fidedigno e válido para avaliar qualidade de vida em idosos.

Outros instrumentos que devido ao fato de serem genéricos e construídos dentro de uma perspectiva transcultural, mostraram-se válidos para o uso nessa população de interesse são o WHOQOL-100 (Moraes, 2004) e WHOQOL-BREVE (Pereira et al. 2006).

Lopes e Rabelo (2006) apresentaram um panorama geral de artigos nacionais e internacionais sobre qualidade de vida e velhice, levantados nos sites científicos AgeLine e Scielo. No Scielo foram encontrados 19 artigos brasileiros sobre qualidade de vida na velhice, e no AgeLine a busca resultou em 553 artigos. Segundo os autores, os resultados dessa investigação corroboraram o reconhecimento de que o construto qualidade de vida é multidimensional e está em crescimento e desenvolvimento no cenário mundial. Este reconhecimento é útil à intervenção voltada para os benefícios individuais e sociais que podem advir da melhoria de qualidade de vida (Lopes e Rabelo 2006).

3.5 Teoria da Resposta ao Item na Avaliação de Qualidade de Vida

A TRI, também chamada de teoria psicométrica moderna, vem ganhando destaque na avaliação de qualidade de vida relacionada à saúde (HRQOL) (Metz et al. 2006; Cella e Chang 2000; Svensson 2001; de Vet et al. 2006; Schünemann et al. 2006; Tennant et al. 2006; Lin 2006).

Há vários questionários que procuram avaliar a qualidade de vida relacionada à saúde, como por exemplo, ADL, que trata das atividades diárias relacionadas à saúde em geral, ESSI, um instrumento com 7 itens relacionados ao suporte social, EORTC QLQ C30, um instrumento com 30 itens para estudos de câncer, FACT-An, uma sub-escala de 13 itens do FACT para anemia, HAQ, um questionário de avaliação de saúde usado em estudos de artrite, PAR-Q, um questionário relacionado às atividades físicas, SDS, uma escala de 18 itens para sintomas de aflição, SF-36 (*Medical Outcomes Study Short-Form*), um questionário de 36 itens relacionados à saúde em geral, WHOQOL-100 e WHOQOL-BREVE, ambos desenvolvidos pela Organização Mundial de Saúde para avaliação de qualidade de vida, este último, uma versão abreviada do anterior com 100 itens, dentre outros (Mesbach et al. 2002).

Um dos primeiros estudos relacionando TRI e HRQOL foi um estudo comparativo entre os modelos de TRI, no caso, o modelo de Rasch e o método

tradicional de soma de escores, aplicados em uma escala de funcionamento físico com 10 itens (PF-10), proveniente do SF-36 (Mielenz et al. 2006), utilizando-se a precisão relativa (Jenkinson et al. 2001). Outro exemplo de avaliação de qualidade de vida relacionada à saúde usando o SF-36 é proveniente do estudo de Huang et al. (2006) onde foram comparadas as propriedades psicométricas e estruturas de fatores do SF-36 e WHOQOL-BREVE, em uma amostra da população de Taiwan, com o objetivo de verificar se os instrumentos mediam os mesmos construtos. O WHOQOL-BREVE, com 26 itens originalmente, foi acrescido de 2 itens próprios para a cultura e costumes locais. Os pesquisadores chegaram à conclusão que os instrumentos não mediam os mesmos construtos, enquanto o SF-36 mede a qualidade de vida relacionada à saúde, o WHOQOL-BREVE mede a qualidade de vida global (O'Carroll et al. 2000).

Em 2004, Noerholm et al. examinaram a validade de construto do WHOQOL-BREVE e chegaram à conclusão de que cada um dos 4 domínios ajustam o modelo logístico de 2 parâmetros, mas o modelo de Rasch não, e que devido à multidimensionalidade, o escore total de 26 itens não ajustou nenhum modelo (Noerholm et al. 2004).

Vários trabalhos salientam a importância da TRI como ferramenta para solucionar problemas que ocorrem utilizando-se os métodos clássicos de análise. Hays et al. (2000) fizeram uma revisão sobre TRI e os desfechos em saúde no século XXI e, com base em um instrumento numa escala de funcionamento físico de 9 itens aplicado aos pacientes do HCSUS (*HIV Cost and Services Utilization Study*), relataram as vantagens potenciais no uso da TRI em medir desfechos em saúde. Em 2003, Prieto et al. a partir do instrumento NHP (*Nottingham Health Profile*) com 38 itens, usando TRI, chegaram a duas versões reduzidas, uma com 20 e outra com 22 itens, ambas apresentando características similares com o original. O instrumento original trata-se de uma medida genérica do status de saúde subjetivo, desenvolvido nos anos 70 na Grã-Bretanha e extensivamente usado na Europa. Contém 38 itens em um formato dicotômico de resposta, descrevendo problemas em 6 dimensões: energia, dor, relação emocional, sono, isolamento social e mobilidade física.

Uma grande contribuição da TRI é a criação e utilização de bancos de itens e testes adaptativos computadorizados (CAT), onde são usados modelos

probabilísticos de medidas para guiar a seleção de questões subseqüentes de um grande banco de itens, maximizando a precisão com um número mínimo de questões. A aplicação médica do CAT (Ware et al. 2000; Cook et al. 2005) é uma inovação recente que permite a simplificação e padronização de desfechos do paciente, monitorando via computador, de um modo útil, prático e válido psicometricamente. Um dos exemplos dessa aplicação é a avaliação do impacto combinado do HIV e seu tratamento sobre o funcionamento e bem estar na perspectiva do paciente, utilizando TRI, banco de itens e CAT's. Como um dos resultados principais chegaram à construção de escalas menores e mais eficientes para medir os múltiplos domínios da qualidade de vida relacionada à saúde de estudos clínicos e estudos observacionais longitudinais (Revicki et al. 1997). Outra aplicação dos CAT's foi realizada por McHorney e Cohen, que desenvolveram um banco de itens para funcionamento físico e equalizaram esses itens usando TRI, a partir de uma amostra de indivíduos idosos. Como resultado final foi gerada uma escala de atividades diárias com 6 itens. Os itens relacionados a se vestir foram os mais discriminantes, seguidos por tomar banho, ir ao banheiro, mobilidade, cozinhar/comer, e atividades comunitárias (McHorney et al. 1997).

Bjorner et al. (2003)^a avaliaram o impacto da dor de cabeça e enxaqueca, utilizando dados de entrevistas do NSHI – *National Survey of Headache Impact* e TRI através do modelo de crédito parcial generalizado. O instrumento utilizado foi o HIT (*Headache Impact Test*) (Bjorner et al. 2003)^b através de testes adaptativos computadorizados (CAT). Adicionalmente foi desenvolvida uma versão para ser auto-administrada, com 6 itens, HIT-6 (Kosinski et al. 2003)^a. Foram observadas vantagens teóricas dos modelos de TRI para avaliar dor de cabeça, além do fato da escala com 5 itens ter alta concordância com o escore baseado em todos os itens. Outro instrumento que mensura o impacto da dor de cabeça e enxaqueca, MSQ (*Migraine Specific Quality of Life Questionnaire*), foi usado em três ensaios clínicos de tratamento para enxaqueca. Foram usados cinco métodos de quantificar a responsividade para estimar e comparar as mudanças do pré para o pós-tratamento, e os resultados exaltam a superioridade da TRI em comparação com os métodos clássicos (Kosinski et al. 2003^b).

Um instrumento para avaliar a fadiga, é o FACIT-F (*Functional Assessment of Chronic Illness Therapy Fatigue Subscale*) que utiliza escalas com 5 alternativas. Esse questionário foi aplicado em 2 grupos, um grupo de pacientes com câncer e o outro da população em geral, utilizando testes adaptativos computadorizados com 13 itens em uma escala de 5 pontos. Para avaliar a fadiga foi usada uma extensão politômica do modelo dicotômico de Rasch (Lai et al. 2003).

Bjorner et. al. (2004) utilizaram o EORTC QLQ-C30 (*European Organization for Research and Treatment of Cancer Quality of Life Group*), desenvolvendo uma versão abreviada com 4 itens para uma escala de funcionamento emocional, apropriado para pesquisas em cuidado paliativo. As escalas de 2 e 3 itens também tiveram uma excelente concordância e pouco viés. Este mesmo instrumento, EORTC QLQ – C30, foi usado no trabalho de Tejada e Rojas (2005), utilizando o modelo de escala gradual em uma medida de qualidade de vida de pacientes terminais de câncer, e no trabalho de Petersen et al. (2005), onde foi apresentado um estudo comparativo entre os modelos de TRI e o método tradicional de soma de escores, através da validade relativa. Neste último trabalho foram analisadas três escalas maiores: funcionamento físico, emocional e fadiga e outras 6 escalas menores. Utilizou-se o modelo de crédito parcial generalizado. Também em 2005, Pagano e Gotay avaliaram a qualidade de vida de 4 grupos étnicos (caucasianos, filipinos, havaianos e japoneses) de pacientes com câncer usando também o instrumento EORTC QLQ C-30, ao qual foi ajustado o modelo de Rasch (Pagano e Gotay 2005). Nas análises foi usada a técnica de DIF (*Differential Item Functioning*) (Kim et al. 2002), outra grande contribuição da TRI, para comparar os diferentes comportamentos e/ou características dos grupos. Mesmo que as análises tenham mostrado um comportamento diferente dos filipinos em auto-relatar e medir sua qualidade de vida, não houve alteração nas propriedades psicométricas do questionário.

Outro instrumento com boas propriedades psicométricas é o MacDQoL com 26 itens, instrumento individualizado para medir o impacto da degeneração macular na qualidade de vida. Usando procedimentos tradicionais foi gerada uma versão reduzida com 22 itens (MacDQoL-22), com excelente consistência interna e uma estrutura de um único fator, podendo ser usado em ensaios clínicos e rotinas de

cuidado clínico (Mitchell et al. 2005). Outro instrumento usado em pacientes com degeneração macular é o DLTV (*Daily Living Tasks Dependent in Vision*) (Hart et al. 2005). Trata-se de um instrumento com 22 itens, em um escala de 4 pontos ordinais, próprio para idosos, visto que a degeneração macular tende a piorar com a idade. Os autores comprovaram a eficiência da TRI em lidar com tal característica, e adicionalmente compararam esses resultados com os da análise fatorial. Não houve diferença significativa nos resultados.

Pallant et al. (2006) desenvolveram e conduziram uma validação preliminar do PIPP (*Perceived Impact of Problem Profile*), instrumento de 23 itens divididos em 5 domínios: auto-cuidado, mobilidade, participação, relacionamentos e bem-estar psicológico, aplicado em uma amostra de 169 adultos com dificuldade de mobilidade. Foi utilizado o modelo de Rasch. São necessárias validações adicionais de seu uso através de diferentes condições de saúde e variantes cenários culturais. O modelo de Rasch também foi utilizado por Wolfe e Kong (1999), desta vez para analisar o WOMAC (*Western Ontario MacMaster Questionnaire*) em 2205 pacientes com osteoartrite, artrite reumatóide e fibromialgia, a fim de verificar o ajuste ao modelo de Rasch, particularmente em respeito à unidimensionalidade. Kopec et al. (2006) também avaliaram uma amostra de pessoas com artrite (n=888) do Canadá. Foram desenvolvidos bancos de itens para 5 domínios de qualidade de vida relacionada à saúde relevantes à artrite (atividades diárias, caminhadas, segurar objetos, dor ou desconforto e sentimentos). Foi utilizado o modelo de crédito parcial generalizado para análise dos itens e análise fatorial para testar unidimensionalidade. Análises adicionais são necessárias para estimar sua performance sob condições dos testes adaptativos computadorizados (CAT), pois sua aplicação na saúde ainda é recente, apesar de comprovadas suas propriedades psicométricas superiores, como validade e fidedignidade.

Um instrumento bastante utilizado é o ALDS (*AMC Linear Disability Score*), referente à capacidade funcional, que tem alcançado propriedades psicométricas para os respondentes em cuidado residencial e forma uma base estável para medir status funcional em uma variedade de situações, incluindo a implementação de CAT's. Foi utilizado o modelo logístico de 2 parâmetros para construir um instrumento com boas

qualidades psicométricas, sendo retirados da versão final os itens que apresentassem estimativas inconsistentes para os parâmetros (Holman et al. 2004).

Holman et al., em 2004 e 2005, também avaliaram o ALDS, primeiramente em idosos e depois em uma população mista, através do modelo logístico de 2 parâmetros, visando a construção de um banco de itens através da calibração na TRI. Os autores atentam para o possível comportamento diferente para outras populações e/ou outros pacientes nesse banco de itens, desenvolvido para quantificar status funcional, expresso pela habilidade de realizar as atividades cotidianas (Holman et al. 2004; Holman et al. 2005).

E Jenkinson et al. (2003) criaram 2 instrumentos para avaliar status de saúde em pacientes com esclerose lateral amiotrófica e outras doenças neuromotores: ALSAQ-40 e ALSAQ-5, uma versão abreviada do anterior, ambos com boas características psicométricas. Foi utilizado o modelo de Rasch na verificação da unidimensionalidade do instrumento original, e com base nos resultados da TRI, foi criada a escala reduzida, com 5 itens, visto que um aspecto fundamental da medida de saúde é a simplicidade de apresentar seus resultados.

O uso correto e crescente da TRI, em especial os CAT's, em lidar com desfechos clínicos relevantes, pode trazer resultados mais precisos, flexíveis e eficientes que os procedimentos clássicos, muitas vezes preferido pela maior simplicidade e/ou falta de conhecimento da TRI. A TRI dá entusiasmo para o prospecto de derivar melhores definições para os construtos subjacentes, e conduz a uma nova esperança para o prospecto de diagnóstico individual. Há um aumento substancial no emprego dos modelos de TRI na avaliação de qualidade de vida, dada sua flexibilidade e eficiência, ainda que em alguns casos, os resultados não diferem dos métodos tradicionais, muito mais simples e difundidos.

4. OBJETIVOS

A avaliação de qualidade de vida não tem sido realizada de forma eficiente e consistente. Os modelos de Teoria da Resposta ao Item, inicialmente empregados na avaliação educacional, vêm ganhando destaque crescente na avaliação de qualidade de vida, e constituem-se em ferramentas promissoras, agregando e utilizando melhor as informações contidas nos instrumentos/questionários de avaliação de qualidade de vida.

Este trabalho tem como objetivo principal estruturar a Teoria da Resposta ao Item, visando comparar o desempenho do modelo de Teoria da Resposta ao Item com os resultados dos métodos tradicionais de análise, dentre eles, a Análise Clássica e a Análise Fatorial, na avaliação de dados produzidos pela Escala de Qualidade de Vida de Flanagan (EQVF) e pelo WHOQOL-100, instrumento elaborado pela Organização Mundial de Saúde em idosos socialmente ativos da região metropolitana de Porto Alegre, RS. Adicionalmente, pretende-se avaliar a equivalência dos resultados para os instrumentos elaborados pela OMS, WHOQOL-100 e WHOQOL-BREVE, este último uma versão abreviada do anterior. O WHOQOL-BREVE foi construído com base nas respostas do instrumento com 100 itens, não sendo aplicado diretamente na amostra disponível.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- Albanese MT. Latent variable models for binary response data. [Tese de Doutorado]. Londres (ENG). The London School of Economics and Political Science; 1990.
- Albanese MT, Camey SA, Marcondes Filho D. Análise das questões que compõem as provas do Concurso Vestibular 93 – UFRGS, usando o Modelo de Variáveis Latentes Logit-Probit. Manuscrito. Porto Alegre; 1993.
- Alexandre JWC, Andrade DF, Vasconcelos AP, Araújo AMS. Uma proposta de análise de um construto para medição dos fatores críticos da gestão pela qualidade por intermédio da Teoria de Resposta ao Item. *Gestão e Produção* 2002; 9(2): 29-141.
- Andrade DF, Tavares HR, Valle RC. Teoria de Resposta do Item: Conceitos e Aplicações. São Paulo: Associação Brasileira de Estatística; 2000.
- Baker FB. The basics of Item Response Theory. New York: ERIC Clearinghouse on Assessment and Evaluation; 2001.
- Bartholomew DJ. Latent variable models and factor analysis. New York: Oxford University Press; 1987.
- Bechger TM, Maris G, Verstralen HHFM, Béguin AA. Using Classical Test Theory in combination with Item Response Theory. *Applied Psychological Measurement* 2003; 27(5): 139-334.
- Berlim MT, Fleck MPA. “Quality of Life”: a brand new concept for research and practice in psychiatry. *Revista Brasileira de Psiquiatria* 2003; 25(4): 249-252.
- Bjorner JB, Kosinski M, Ware JE Jr. Using item response theory to calibrate the Headache Impact Test (HIT) to the metric of traditional headache scales. *Quality of Life Research* 2003; 12(8): 981-1002.
- Bjorner JB, Petersen MA, Groenvold M, Aaronson N, hlnner-Elmqvist M, Arraras JI, et al. Use of item response theory to develop a shortened version of the EORTC QLQ-C30 emotional functioning scale. *Quality of Life Research* 2004; 13(10): 1683-1697.
- Bjorner JB, Kosinski M, Ware JE Jr. Calibration of an item pool for assessing the burden of headaches: An application of Item Response Theory to the Headache Impact Test (HITTM). *Quality of Life Research* 2003; 12: 913-933.
- Cassiano JG, Dias JMD, Teixeira-Salmela LC, Pereira GS, Magalhães CM. Promovendo Saúde e Qualidade de Vida em Adultos Maduros e Idosos. Anais do 8º Encontro de Extensão da UFMG, Belo Horizonte, 2005.

Cella D, Chang C-H. A discussion of Item Response Theory and its application in health status assessment. *Medical Care* 2000; 38(2 suppl II): II66-II72.

Cella D, Gershon R, Lai J-S, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Quality of Life Research* 2007 (online).

Chaimovicz F. A saúde dos idosos brasileiros às vésperas do século XXI: problemas, projeções e alternativas. *Revista de Saúde Pública* 1997; 31(2): 184-200.

Cook KF, O'Malley KJ, Roddey TS. Dynamic Assessment of Health Outcomes: Time to Let the CAT Out of the Bag? *Health Research and Educational Trust* 2005; 40(5): 1649-1711.

Crane PK, Gibbons LE, Narasimhalu K, Lai J-S, Cella D. Rapid detection of differential item functioning in assessments of health-related quality of life: The Functional Assessment of Cancer Therapy. *Quality of Life Research* 2007; 16(1): 101-114.

Cruz LN. Associação entre qualidade de vida e depressão em pacientes portadores de doenças crônicas. [Dissertação de Mestrado]. Porto Alegre (RS). Programa de Pós-graduação em Epidemiologia, Faculdade de Medicina da Universidade Federal do Rio Grande do Sul; 2005.

Cúri M. Análise de questionários com itens constrangedores [Tese de Doutorado]. São Paulo (SP). Instituto de Matemática e Estatística, Universidade de São Paulo; 2006.

de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health and Quality of Life Outcomes* 2006; 4(1): 54-59.

Embretson SE, Reise. *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum Associates; 2000.

Farquhar M. Elderly people's definitions of quality of life. *Social Science and Medicine* 1995; 41(10):1439-1446.

Fleck MPA. O instrumento de avaliação de qualidade de vida da Organização Mundial de Saúde (WHOQOL-100): características e perspectivas. *Ciência & Saúde Coletiva* 2000; 5(1): 33-38.

Fleck MPA, Borges ZN, Bolognesi G, Rocha NS. Desenvolvimento do WHOQOL, módulo espiritualidade, religiosidade e crenças pessoais. *Revista de Saúde Pública* 2003; 37(4): 446-455.

Fleck MPA, Chachamovich E, Trentini CM. Projeto WHOQOL-OLD: método e resultado de grupos focais no Brasil. *Revista de Saúde Pública* 2003; 37(6): 793-799.

Fleck MPA, Chachamovich E, Trentini C. Development and validation of the Portuguese version of the WHOQOL-OLD module. *Revista de Saúde Pública* 2006; 40(5): 785-791.

Fleck MPA, Leal OF, Louzada S., Xavier M., Chachamovich E., Vieira G., et al. Desenvolvimento da versão em português do instrumento de avaliação de qualidade de vida da OMS (WHOQOL – 100). *Revista Brasileira de Psiquiatria* 1999; 21(1): 19-28.

Fleck MPA, Leal OF, Louzada S., Xavier M., Chachamovich E., Vieira G., Santos L, et al. Aplicação da versão em português do instrumento abreviado de avaliação da qualidade de vida “WHOQOL – Bref”. *Revista de Saúde Pública* 2000; 34(2): 178-183.

Hambleton RK. Emergence of Item Response Modelling in Instrument Development and Data Analysis. *Medical Care* 2000; 38(9): 60-65.

Hart PM, Stevenson MR, Montgomery A-H, Muldrew KA, Chakravarthy U. Further validation of the Daily Living Tasks Dependent on Vision: identification of domains. *British Journal of Ophthalmology* 2005; 89: 1127-1130.

Hays RD, Lipscomb J. Next steps for use of item response theory in the assessment of health outcomes. *Quality of Life Research* 2007. <http://www.springerlink.com/w10048.dotlib.com.br/content/pp368578887q3031/?p=783b07f51fec46698312c2fd554f278b&pi=0> (acessado em 20 de abril de 2007).

Hays R, Morales L, Reise SP. Item Response Theory and Health Outcomes Measurement in the 21st century. *Medical Care* 2000; 38(9 suppl II): II28-II42.

Holland PW, Hoskens M. Classical test theory as a first-order Item Response Theory: Application of true score prediction from a possibly non-parallel test. *Psychometrika* 2003; 68(1): 123-149.

Holman R, Lindeboom R, Vermeulen M, Haan RJ. The AMC Linear Disability Score Project in a population requiring residential care: psychometric properties. *Health and Quality of Life Outcomes* 2004; 2: 42.

Holman R, Glas CAW, Lindeboom R, Zwindermann AH, Haan RJ. Practical methods for dealing with ‘not applicable’ item responses in the AMC Linear Disability Score Project. *Health and Quality of Life Outcomes* 2004; 2: 29.

Holman R, Weisscher N, Glas CAW, Dijkgraaf MGW, Vermeulen M, Haan RJ, Lindeboom R. The Academic Medical Center Linear Disability Score (ALDS) item bank: item response theory analysis in a mixed patient population. *Health and Quality of Life Outcomes* 2005; 3: 83.

Huang C, Wu AW, Frangakis C. Do the SF-36 and WHOQOL-Bref measure the same constructs? Evidence from the Taiwan population. *Quality of Life Research* 2006; 15: 15-24.

Iezzoni LI. Going beyond disease to address disability. *New England Journal of Medicine* 2006; 355 (10): 976-979.

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). *Journal of Neurology, Neurosurgery and Psychiatric* 2001; 71: 220-224.

Jenkinson C, Norquist JM, Fitzpatrick R. Deriving summary indices of health status from the Amyotrophic Lateral Sclerosis Assessment Questionnaires (ALSAQ-40 and ALSAQ-5). *Journal of Neurology, Neurosurgery and Psychiatric* 2003; 74: 242-245.

Kalache A, Veras RP, Ramos LR. O envelhecimento da população mundial: um desafio novo. *Revista de Saúde Pública* 1987; 21(3): 200-210.

Kim Y, Pilkonis PA, Frank E, Thase ME, Reynolds CF. Differential Functioning of the Beck Depression Inventory in Late life Patients: Use of Item Response Theory. *Psychology and Aging* 2002; 17(3): 179-191.

Kopec JA, Sayre EC, Davis AM, Badley EM, Abrahamowicz M, Sherlock L, Williams JI, Anis AH, Esdaile JM. Assessment of health-related quality of life in arthritis: conceptualization and development of five item banks using item response theory. *Health and Quality of Life Outcomes* 2006; 4:33.

Kosinski M, Bjorner JB, Ware JE Jr, Batenhorst A, Cady RK. The responsiveness of headache impact scales scored using 'classical' and 'modern' psychometric methods: a re-analysis of three clinical trials. *Quality of Life Research* 2003; 12: 903-912.

Kosinski M, Bayliss Ms, Bjorner JB, Ware JE Jr, Garber WH, Batenhorst A, et al. a six item short-form survey for measuring headache impact: The HIT-6TM. *Quality of Life Research* 2003; 12: 963-974.

Lai JS, Cella D, Chang CH, Bode RK, Heinemann AW. Item banking to improve, shorten and computerize self-reported fatigue: an illustration of steps to create a core item bank from the FACIT-Fatigue Scale. *Quality of Life Research* 2003; 12(5): 485-501.

Lai J-S, Crane PK, Cella D. Factor analysis techniques for assessing sufficiency of unidimensionality of cancer related fatigue. *Quality of Life Research* 2006; 15: 1179-1190.

Lentz RS, Costenaso RGS, Gonçalves LHT, Nassar SM. O profissional de enfermagem e a qualidade de vida: uma abordagem fundamentada nas dimensões propostas por Flanagan. *Revista Latino-americana de Enfermagem* 2000; 8(4): 7-14.

Lin TH. Identifying optimal items in quality of life assessment. *Quality & Quantify* 2006. <http://www.springerlink.com.w10048.dotlib.com.br/content/e4t8455659871078/?p=eaf19e5e052a4d4b974080350e0087bb&pi=0> (acessado em 12 de dezembro de 2006)

Linden WJ, Hambleton RK. *Handbook of Modern Item Response Theory*. New York: Springer-Verlag; 1997.

Lohr KN. Assessing health status and quality of life instruments: Attributes and review criteria. *Quality of Life Research* 2002; 11: 193-205.

Lopes EVL, Rabelo DF. Qualidade de vida na velhice: indicadores nacionais e internacionais. *Kairós* 2006; 9(2): 209-226.

McHorney CA, Haley SM, Ware JE Jr. Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. *Journal of Clinical Epidemiology* 1997; 50(4): 451-461.

McHorney CA, Cohen AS. Equating Health Status Measures with Item Response Theory. *Medical Care* 2000; 38 (9): 43-59.

McHoney CA. Ten recommendations for advancing patient-centered outcomes measurement for older persons. *Annals of Internal Medicine* 2003; 139(5): 403-409.

Mesbach M, Cole BF, Lee M-LT. *Statistical Methods for Quality of Life Studies*. London: Kluwer Academic Publishers; 2002.

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. *Quality of Life Research* 2006; 15: 1639-1649.

Mielenz T, Jackson E, Currey S, DeVellis R, Callahan LF. Psychometric properties of the Centers for Disease Control and Prevention Health-Related Quality of Life (CDC HRQOL) items in adults with arthritis. *Health and Quality of Life Outcomes* 2006; 4:66.

Minayo MCS, Hartz ZMA, Buss PM. Qualidade de vida e saúde: um debate necessário. *Ciência & Saúde Coletiva* 2000; 5(1): 7-18.

Mitchell J, Wolffsohn JS, Woodcock A, Andersen SJ, McMillan CV, ffytche T, et al. Psychometric evaluation of the MacDQoL individualized measure of the impact of macular degeneration on quality of life. *Health and Quality of Life Outcomes* 2005; 3: 25.

Moraes JF. Fatores determinantes do envelhecimento bem-sucedido do idoso socialmente ativo da região metropolitana de Porto Alegre [Tese de Doutorado]. Porto Alegre (RS). Programa de Pós-graduação em Gerontologia Médica, Pontifícia

Universidade Católica do Rio Grande do Sul; 2004.

Noerholm V, Groenvold M, Watt T, Bjorner JB, Rasmussen NA, Bech P. Quality of life in the Danish general population-normative data and validity of WHOQOL-Bref using Rasch and item response theory models. *Quality of Life Research* 2004; 13: 531-540.

O'Carroll, Smith K, Coustan M, Cossar JA, Hayes Pc. A comparison of the WHOQOL-BREF in detecting change in quality of life following liver transplantation. *Quality of Life Research* 2000; 9: 121-124.

Pagano IS, Gotay CC. Ethnic differential item functioning in the assessment of quality of life in cancer patients. *Health and Quality of Life Outcomes* 2005; 3:60.

Paixão CM Jr, Reichenheim ME. Uma revisão sobre instrumentos de avaliação do estado funcional do idoso. *Cadernos de Saúde Pública* 2005; 21(1): 7-19.

Pallant JF, Misajon RA, Bennett E, Mandersen L. Measuring the impact and distress of health problems from the individual's perspective: development of the Perceived Impact of Problem Profile (PIPP). *Health and Quality of Life Outcomes* 2006; 4: 36.

Paschoal SMP. Qualidade de vida do idoso: elaboração de um instrumento que privilegia sua opinião [Dissertação de Mestrado]. São Paulo (SP). Faculdade de Medicina, Universidade de São Paulo; 2000.

Pasquali L, Primi R. Fundamentos da Teoria de Resposta ao Item – TRI. *Avaliação Psicológica* 2003; 2(2): 99-110.

Pereira RJ, Cotta RMM, Franceschini SCC, Ribeiro RCL, Sampaio RF, Priore SE, Cecon PR. Contribuição dos domínios físico, social, psicológico e ambiental para a qualidade de vida global de idosos. *Revista de Psiquiatria do Rio Grande do Sul* 2006; 28(1): 27-38.

Petersen MA, Groenvold M, Aaronson N, Brenne E, Fayers P, Nielsen JD, et al. Scoring based on item response theory did not alter the measurement ability of EORTC QLQ-C30 scales. *Journal of Clinical Epidemiology* 2005; 58(9): 902-908.

Presas DS. A teoria de resposta ao item na avaliação educacional: Uma análise das questões que compõem a prova de matemática do concurso vestibular 2004 – UFRGS [Monografia]. Porto Alegre (RS). Departamento de Estatística, Universidade Federal do Rio Grande do Sul; 2005.

Prieto L, Alonso J, Lamarca R. Classical test theory versus Rasch analysis for quality of life questionnaire reduction. *Health and Quality of Life Outcomes* 2003; 1: 27.

Ramos LR, Veras RP, Kalache A. Envelhecimento populacional: uma realidade brasileira. *Revista de Saúde Pública* 1987; 21(3): 211-224.

Ramos LR. Fatores determinantes do envelhecimento saudável em idosos residentes em centro urbano: Projeto Epidoso, São Paulo. *Cadernos de Saúde Pública* 2003; 19(3): 793-798.

Revicki DA, Cella DF. Health status assessment for the twenty-first century: item response theory, item banking and computer adaptive testing. *Quality of Life Research* 1997; 6(6): 595-600.

Santos SR, Santos IBC, Fernandes MGM, Henrique MERM. Qualidade de vida do idoso na comunidade: aplicação da Escala de Flanagan. *Revista Latino-americana de Enfermagem* 2002; 10(6): 757-764.

Schünemann HJ, Akl EA, Guyatt GH. Interpreting the results of patient reported outcomes measures in clinical trials: the clinician's perspective. *Health and Quality of Life Outcomes* 2006; 4(1): 62-69.

Soares TM. Utilização da Teoria de Resposta ao Item na produção de indicadores sócio-econômicos. *Pesquisa Operacional* 2005; 25(1): 83-112.

Sousa L, Galante H, Figueiredo D. Qualidade de vida e bem-estar de idosos: um estudo exploratório na população portuguesa. *Revista de Saúde Pública* 2003; 37:364-371.

Svensson E. Construction of a single global scale for multi-item assessment of the same variable. *Statistics in Medicine* 2001; 20: 3831-3846.

Takane Y, Leeuw J, On the relationship between Item Response Theory and Factor Analysis of Discretized Variables. *Psychometrika* 1987; 52(3): 393-408.

Tejada AJR, Rojas OML. Application of an IRT polytomous model for measuring Health Related Quality of Life. *Social Indicators Research* 2005; 74(2): 369-394.

Tennant AH, Küçükdeveci AA, Kertlay S, Elhan AH. Assessing normative cut points through differential item functioning analysis: An example from the adaptation of the Middlesex Elderly Assessment of the Mental State (MEAMS) for use as a cognitive screening test in Turkey. *Health and Quality of Life Outcomes* 2006; 4(1): 18-25.

Valle RC. Teoria de Resposta ao Item [Dissertação de Mestrado]. São Paulo (SP). Instituto de Matemática e Estatística, Universidade de São Paulo; 1999.

van Belle G, Arnold A. Reliability of cognitive tests used in Alzheimer's disease. *Statistics in Medicine* 2000; 19(11-12): 1411-1420.

Vecchia RD, Ruiz T, Bocchi SCM, Corrente JE. Qualidade de vida na terceira idade: um conceito subjetivo. *Revista Brasileira de Epidemiologia* 2005; 8(3): 246-252.

Vendramini CMM, Silva MC, Canale M. Análise de itens de uma prova de raciocínio estatístico. *Psicologia em Estudo* 2004; 9(3): 487-498.

Veras RP. Em busca de uma assistência adequada à saúde do idoso: revisão da literatura e aplicação de um instrumento de detecção precoce e de previsibilidade de agravos. *Cadernos de Saúde Pública* 2003; 19(3): 705-715.

Veras RP, Ramos LR, Kalache A. Crescimento da população idosa no Brasil: transformações e conseqüências na sociedade. *Revista de Saúde Pública* 1987; 21(3): 225-233.

Vergara LGL. Avaliação do Ensino de Ergonomia para o Design aplicando a Teoria de Resposta ao Item (TRI) [Tese de Doutorado]. Florianópolis (SC) Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina; 2005.

Vianna HM. Testes em Educação. São Paulo: IBRASA – Instituição Brasileira de Difusão Cultural S.A; 1978.

Vitória F, Almeida LS, Primi R. Unidimensionalidade em testes psicológicos: conceito, estratégias e dificuldades na sua avaliação. *Psic* 2006; 7(1): 1-7.

Ware JE Jr, Bjorner JB, Kosinski M. Practical implications of Item Response Theory and Computerized Adaptive Testing. *Medical Care* 2000, 38 (suppl II): II43-II59.

Ware JE Jr. Conceptualization and measurement of health-related quality of life: comments on an evolving field. *Archives of Physical Medicine and Rehabilitation* 2003; 84(2): 43-51.

Wolfe F, Kong SX. Rasch analysis of the Western Ontario MacMaster Questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Annals of Rheumatic Diseases* 1999; 58: 563-568.

Xavier FMF, Ferraz MPT, Marc N, Escosteguy NU, Moriguchi EH. Elderly people's definition of quality of life. *Revista Brasileira de Psiquiatria* 2003; 25(1): 31-39.

6. ARTIGO

Comparação do desempenho da Teoria da Resposta ao Item com métodos tradicionais na avaliação de instrumentos de qualidade de vida aplicados a idosos

Performance's comparison of Item Response Theory with traditional methods
in the evaluation of quality of life instruments in elderly

Katrine Guewehr ¹

João Riboldi ^{2,3}

1. Mestranda do Programa de Pós-Graduação em Epidemiologia, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul

2. Programa de Pós-Graduação em Epidemiologia, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul

3. Departamento de Estatística, Universidade Federal do Rio Grande do Sul

Correspondência

Katrine Guewehr

Rua Carlos Reinaldo Muller, 14, bairro Liberdade, Novo Hamburgo, RS, Brasil

Telefone: +55 51 9601-3844

Email: katrine_guewehr@yahoo.com.br

A ser enviado: Cadernos de Saúde Pública

Resumo

O objetivo deste estudo foi a comparação do desempenho da Teoria da Resposta ao Item (TRI) com os métodos tradicionais de análise, Análise Fatorial (AF) e Teoria Clássica dos Testes (TCT), para os instrumentos aplicados em idosos, WHOQOL-100 e EQVF, além do WHOQOL-BREVE, construído com base nas respostas ao instrumento maior. Foi utilizado na TRI o Modelo Logístico de 2 Parâmetros para a estimação da qualidade de vida global e por domínios a partir dos resultados obtidos para cada um dos instrumentos. As três escalas apresentaram valores altos de fidedignidade. Com base na aplicação do modelo de TRI na EQVF, pôde-se concluir que os idosos tiveram maior dificuldade de sentirem-se satisfeitos com sua aprendizagem, e que indivíduos com um bom nível de autoconhecimento, trabalho, boa comunicação e oportunidades de lazer apresentaram um nível superior de qualidade de vida. Através do WHOQOL-BREVE, verificou-se que estes idosos relataram maior dificuldade de estarem satisfeitos com sua qualidade de vida em geral, e deram grande importância para a capacidade de desempenhar as atividades cotidianas. No WHOQOL-100, foram encontrados baixos níveis de satisfação com a memória, vida sexual e situação financeira. Os resultados via TRI não diferem significativamente dos resultados encontrados pela Teoria Clássica dos Testes e pela Análise Fatorial, muito embora potencializem favoravelmente o uso da TRI na avaliação de qualidade de vida de idosos.

Palavras-chaves: qualidade de vida, escalas, idosos, análise fatorial.

Abstract

The objective of this research was the performance's comparison of Item Response Theory with traditional methods, Factor Analysis and Classical Theory Test, for the instruments applied in elderly, WHOQOL-100 and FQLS, beyond the WHOQOL-BREF, constructed with answers of the WHOQOL-100. Using the 2 parameters logistic model on the estimation of global life's quality as well domains from the obtained results for each of the instruments. The three scales presented high values of reliability. Based on application of IRT model on the EQVF could be concluded that elderly have difficulties on feeling satisfaction with their learning process, and that individuals with good level of self-knowledge, job, good communication and leisure opportunities presented a superior level on life's quality. Through the WHOQOL-BREF was possible verify they presented more difficult with their life's quality in general and really care about the capacity to accomplish their daily activities. With WHOQOL-100 was founded low level with memory, sexual life and financial situation. We conclude that the tests results obtained with IRT aren't significantly different from Classical Test Theory and Factor Analysis, but Item Response Theory was confirmed as a promising to evaluation of quality of life.

Key words: life's quality, scales, elderly, factor analysis.

Introdução

O envelhecimento da população é um fenômeno de amplitude mundial. Atualmente, no Brasil, a população de idosos, definidos pela Organização Mundial de Saúde (OMS) como indivíduos com idades maiores ou iguais a 60 anos, já ultrapassa os 7%, e as previsões indicam que alcançará os 18 milhões em 2010. Levando em conta a continuidade das tendências verificadas para as taxas de fecundidade e longevidade da população brasileira, as estimativas para os próximos 20 anos indicam que a população idosa poderá exceder 30 milhões ao final deste período, chegando a representar aproximadamente 13% da população.¹⁻³ Em 2005, no levantamento da PNAD (Pesquisa Nacional por Amostra de Domicílios), referente ao ano de 2004, a população de idosos no Brasil já se aproximava dos 17 milhões (9,8% da população total). A expectativa de vida também aumentou a partir de 1980, de 62,6 para 71,7 anos.⁴ E este rápido crescimento previsto para a população idosa nas próximas décadas, especialmente no Brasil, indica a necessidade de se estimular o estudo da sociologia da terceira idade no contexto do Terceiro Mundo.^{3,5-7}

A velhice deixou de ser sinônimo de doença, solidão e dependência. Dado o aumento relativo e absoluto de idosos na população, envelhecer com qualidade de vida passou a ser um desafio para todos. A qualidade de vida na terceira idade tem sido, muitas vezes, associada a questões de dependência-autonomia. As dependências observadas nos idosos resultam tanto de alterações biológicas (deficiências ou incapacidades) como de mudanças nas exigências sociais (desvantagens), e freqüentemente, estas últimas parecem determinar as primeiras.⁸⁻¹⁰

Não há um consenso sobre o conceito de qualidade de vida, ainda que exista uma variedade de ferramentas quantitativas e qualitativas usadas para medir e/ou descrever aspectos da qualidade de vida dos idosos, tornando esse tema algo novo.

Entende-se por qualidade de vida, de acordo com a OMS, como "a percepção do indivíduo de sua posição na vida no contexto da cultura e sistema de valores nos quais ele vive e em relação aos seus objetivos, expectativas, padrões e preocupações"¹¹. A qualidade de vida é subjetiva, multidimensional por natureza e como tal, apresenta dimensões positivas e negativas, sendo fundamental avaliar essa

multidimensionalidade conforme sua percepção pelo idoso, ou seja, o quanto está insatisfeito ou satisfeito com a sua qualidade de vida.¹²⁻¹⁷

Aferir qualidade de vida é muito complexo e os domínios da qualidade de vida que requerem medida entre os idosos incluem os problemas de saúde, que podem levar à incapacidade e invalidez, saúde mental, habilidade funcional, estado geral de saúde, satisfação de vida, estado de espírito, controle (autonomia) e suporte social. Isso mostra a enorme variedade de dimensões a serem aferidas numa avaliação de qualidade de vida em idosos, levando à necessidade de instrumentos multidimensionais que possam captar a enorme variabilidade dos diferentes grupos de idosos. Tais instrumentos devem levar em conta as especificidades dessa faixa de idade, como seus valores e experiências de vida, fatores ligados à saúde e diversas situações como aposentadoria, viuvez, dependência, perda de autonomia e de papéis sociais. Todos esses fatores aumentam a complexidade de tal avaliação. Quando se avalia qualidade de vida em idosos, geralmente usa-se um instrumento de uso universal, construído e validado para diversas faixas de idade e populações, como os utilizados nesse presente estudo: WHOQOL-100 (Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS)^{11,18}, WHOQOL-BREVE (Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS – versão abreviada)¹⁹⁻²² e a EQVF (Escala de Qualidade de Vida de Flanagan)^{23,24}. Apenas em 2005, o Grupo WHOQOL, com a colaboração de vários centros internacionais, desenvolveu e validou um instrumento específico para idosos: o WHOQOL-OLD, composto por 24 itens, divididos em 6 facetas. As facetas referem-se ao funcionamento dos sentidos, autonomia, atividades passadas, presentes e futuras, participação social, morte e morrer, e intimidade, contemplando aspectos relevantes não presentes nos instrumentos gerais.²⁵ No Brasil, Fleck e colaboradores desenvolveram e validaram uma versão em português do WHOQOL-OLD, que apresentou boas características psicométricas.²⁶

Por muitas décadas a qualidade de vida, assim como outros desfechos, foi mensurada através dos métodos clássicos de análise²⁷, baseados nas somas dos escores, sendo, portanto, dependente direto do instrumento escolhido. Um indivíduo submetido a diferentes instrumentos, mesmo que com o mesmo objetivo, tinha escores diferentes e não podiam ser comparados seus resultados. Essa técnica, ainda

utilizada hoje em dia em alguns testes psicológicos²⁸, apresenta várias limitações, e a Teoria da Resposta ao Item (TRI), também conhecida como teoria psicométrica moderna, surgiu com o intuito de solucionar vários dos problemas que os métodos clássicos apresentam. A TRI não entra em contradição com a Teoria Clássica dos Testes (TCT), mas traz uma nova proposta de análise, cujo elemento central é o item e não o instrumento como um todo, além de apresentar novos recursos tecnológicos para a avaliação de qualidade de vida, com uma maior consistência e eficiência. A TRI²⁹⁻³⁷ propõe modelos para os traços latentes, que são características não observadas diretamente, sendo inferidas a partir de variáveis secundárias que estejam relacionadas a ela. O traço latente avaliado neste trabalho é a qualidade de vida global dos idosos, sendo estimada a partir das respostas aos itens que constam em cada um dos instrumentos utilizados.

Sendo assim, este estudo visa comparar o desempenho da TRI com os métodos tradicionais de análise (Análise Fatorial e Teoria Clássica dos Testes) na avaliação de qualidade de vida de idosos socialmente ativos da região metropolitana de Porto Alegre, avaliada através da Escala de Qualidade de Vida de Flanagan, WHOQOL-100, e por construção, WHOQOL-BREVE. Paralelamente, busca-se comparar a equivalência dos resultados do WHOQOL-100 com WHOQOL-BREVE.

Métodos

Fonte de Dados

A base de dados utilizada nesse trabalho é proveniente de um estudo observacional realizado por Moraes (2004)¹, com 400 idosos socialmente ativos, com 60 anos ou mais, incluídos por conveniência, que responderam os seguintes instrumentos: Escala de Qualidade de Vida de Flanagan e o WHOQOL-100, além da Escala de Depressão e de Desesperança de Beck³⁸. O instrumento reduzido, WHOQOL-BREVE foi construído a partir do instrumento original com 100 itens (WHOQOL-100).

Delineamento

Estudo transversal com 400 idosos socialmente ativos da região metropolitana de Porto Alegre.

Instrumentos de Avaliação

A qualidade de vida será estimada a partir das respostas aos instrumentos de qualidade de vida aplicados. Outras variáveis sócio-demográficas, como sexo, idade, nível educacional, estado civil, percepções sobre sua saúde atual, também foram coletadas. A Escala de Qualidade de Vida de Flanagan, com 15 itens é dividida em 5 domínios: bem estar físico e mental, relações com outras pessoas, atividades sociais, comunitárias e cívicas, desenvolvimento pessoal e realização e recreação, avaliadas em uma escala de 7 pontos. Já o WHOQOL-100, possui 96 itens divididos em 24 facetas e 6 domínios ou construtos: físico, psicológico, nível de dependência, relação sociais, ambiente e espiritualidade / crenças / religiões, além dos 4 itens de qualidade de vida em geral. E o WHOQOL-BREVE, possui 26 itens, 24 itens inclusos nos 4 domínios (físico, psicológico, nível de independência e meio ambiente) e 2 questões gerais. Ambos têm em sua estrutura original uma escala de resposta de 5 pontos. Também foram consideradas as respostas obtidas para as Escalas de Depressão e Desesperança de Beck.

Estratégia Analítica

Inicialmente os dados (respostas aos itens para os 3 instrumentos) foram dicotomizados, visando uma maior homogeneidade, já que para um grande número de itens as respostas se concentravam geralmente em 2 ou 3 categorias. A EQVF, construída originalmente com 7 categorias de resposta (1 = muito insatisfeito a 7 = muito satisfeito) foi recategorizada, onde as respostas de 1 a 5 assumiram o valor 0 (não satisfeito) e as categorias 6 e 7 assumiram o valor 1 (satisfeito). Já para o WHOQOL-100 e WHOQOL-BREVE, que seguem o mesmo padrão de 5 categorias de resposta, divididas em escalas de intensidade (nada-extremamente), capacidade (nada-completamente), frequência (nunca-sempre) e avaliação (muito insatisfeito-muito satisfeito, muito ruim-muito bom), foram categorizados da seguinte forma

geral: respostas de 1 a 3 receberam o valor 0, que equivale a não satisfeito e respostas 4 e 5 como valor 1, referente a satisfação com determinado item.

O modelo de TRI usado foi o modelo logístico de 2 parâmetros (ML-2), mais adequado à estrutura dos dados, e as análises foram conduzidas em 2 etapas: estimação da qualidade de vida global para cada uma das escalas e a estimação por domínios, também para cada uma das escalas, com exceção da EQVF, de apenas 15 itens, uma vez que não seria possível a estimação por domínio, o que levaria à inconsistência dos parâmetros estimados.

Para um melhor ajuste do ML-2 foram retirados das escalas os itens com correlação bisserial (correlação linear do escore do item com o escore bruto total) muito baixa (inferiores a 0,25), e que apresentassem pequenos valores para as estimativas do parâmetro de inclinação/discriminação (inferiores a 0,4). Esse valor de 0,4 parece razoável, visto que alguns trabalhos utilizam o ponto de corte 0,7, o que levaria à retirada de muitos itens, principalmente no WHOQOL-100, fazendo com que o instrumento perdesse a sua caracterização original. Não há consenso sobre qual o valor mínimo do parâmetro de inclinação ^(21,29,36).

Adicionalmente foram realizadas a análise fatorial exploratória, com método de rotação Varimax e normalização de Kaiser com os itens dicotomizados; análises de variância, onde as variáveis dependentes são o traço latente produzido via TRI para cada escala e o escore bruto das mesmas, e os fatores sexo, idade categorizada, escolaridade, estado civil, saúde percebida, escala de desesperança e de depressão; análises clássicas dos escores brutos e latentes; correlações lineares entre os escores brutos e latentes, tanto globais, como por domínios; estatística de concordância Kappa com as classificações obtidas via mediana dos escores brutos e latentes; e avaliação da consistência interna através do coeficiente alfa de Cronbach. Todas as análises, com exceção da avaliação da fidedignidade das escalas foram realizadas somente com os itens consistentes remanescentes do modelo de TRI, para viabilizar as comparações.

As análises referentes à TRI foram desenvolvidas no BILOG ³⁹. Para a Teoria Clássica das Medidas e Análise Fatorial foram utilizados o SAS (SAS Institute, Cary NC, EUA) e SPSS (SPSS Inc, Chicago, EUA).

Teoria da Resposta ao Item aplicada à Qualidade de Vida

O modelo logístico de 2 parâmetros é adequado para situações em que o construto é medido por meio de um questionário estruturado, isto é, objetivo, com respostas dicotomizadas, ou quando a análise é realizada de forma dicotomizada dentro de uma avaliação subjetiva (aberta).

O modelo logístico de 2 parâmetros é definido por:

$$P(X_{ij} = 1 / \theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \quad (1)$$

com $i = 1, 2, 3, \dots, I$ (representando os I itens referentes aos aspectos de qualidade de vida); e $j = 1, 2, 3, \dots, n$ representando os n idosos inclusos na pesquisa.

Classificando-se como satisfeito ou não com determinado item ou condição avaliada na questão como forma dicotomizada de representar os itens, interpreta-se o modelo (1) como:

X_{ij} , variável dicotômica que assume o valor 1 quando o j -ésimo idoso sente-se satisfeito com o i -ésimo item referente à qualidade de vida, ou assume o valor 0 quando o j -ésimo idoso sente-se insatisfeito com o i -ésimo item.

θ_j : é a habilidade do j -ésimo indivíduo (idoso), isto é, a qualidade de vida latente do idoso inferida a partir da análise dos itens.

$P(X_{ij} = 1 / \theta_j)$ é a probabilidade do j -ésimo idoso com qualidade de vida latente θ_j sentir-se satisfeito com o i -ésimo item da escala de qualidade de vida.

a_i é o parâmetro que representa o poder de discriminação no item i , isto é, discrimina os indivíduos idosos que referem alta qualidade de vida dos que referem baixa qualidade de vida. É proporcional à derivada da tangente da curva no ponto de inflexão, ou seja, à inclinação da curva característica do item.

b_i é o parâmetro que representa o grau de dificuldade do idoso para sentir-se satisfeito no item i , medido na mesma escala da qualidade de vida latente.

D é uma constante (um fator de escala) e é igual a 1. Quando se deseja que a função logística forneça resultados semelhantes ao da função ogiva normal utiliza-se o valor de 1,7.

A relação entre os parâmetros pode ser representada e melhor entendida através da curva característica do item (Figura 1). À medida que os escores de qualidade de vida aumentam, maior a probabilidade do idoso estar satisfeito com o item avaliado.

Estimativa dos parâmetros e recursos computacionais utilizados

A estimação dos parâmetros dos modelos em TRI pode ser procedida de 3 formas: a estimação das habilidades, quando já se conhecem os parâmetros dos itens, a estimação dos parâmetros dos itens quando já se conhecem as habilidades e a estimação conjunta das habilidades e dos parâmetros dos itens, como no presente caso. O processo de estimação dos itens é conhecido como calibração.

Geralmente e para qualquer uma das situações usa-se o método da máxima verossimilhança, usando algum processo iterativo como os algoritmos de Newton-Raphson ou de Scoring de Fisher. Mais recentemente, métodos bayesianos foram propostos para, entre outras coisas, resolver o problema de estimação dos parâmetros dos itens. Neste trabalho optou-se pelo método da máxima verossimilhança para a estimação dos parâmetros dos itens e das habilidades.

O software estatístico utilizado na análise dos dados deste estudo foi o BILOG³⁹, que executa a análise em três etapas (fases 1, 2 e 3), que se caracterizam pelo tipo de tarefas realizadas em cada uma delas. Na fase 1, fase de entrada e leitura de dados, os arquivos de saída fornecem algumas estatísticas descritivas e correlações de interesse, usadas na Teoria Clássica de Medida. A fase 2 é a fase de calibração dos itens. Nessa fase são estimados os parâmetros dos itens. A estimativa das habilidades (qualidade de vida latente) dos respondentes (idosos) é executada na fase 3. O BILOG fornece ainda gráficos contendo algumas informações de interesse, como: curvas características e curvas de informação de cada item e do teste, retas de regressão das habilidades e gráfico de ajuste das habilidades à curva normal.

Interpretação dos parâmetros e adequação do modelo

O parâmetro a_i (discriminação) é proporcional à derivada da tangente da curva no ponto de inflexão, ou seja, à inclinação da curva característica do item (CCI) no ponto b_i (dificuldade), conforme mostra a Figura 1. Por este motivo, a_i

também é chamado de parâmetro de inclinação do item i . É um parâmetro estritamente positivo, pois itens com a_i negativo indicariam que indivíduos com alto nível de satisfação tendem a ter baixa qualidade de vida. Valores muito altos indicam uma CCI muito íngreme, discriminando os indivíduos que possuem qualidade de vida abaixo do parâmetro b_i e os que possuem nível de qualidade de vida acima do valor de b_i .

Uma questão com alto poder de discriminação indica que os idosos com maior qualidade de vida tenderão a sentirem-se mais satisfeitos nesse determinado item. Se a discriminação for baixa, idosos com diferentes níveis de qualidade de vida têm a mesma probabilidade de sentirem-se satisfeitos.

O parâmetro de dificuldade b_i , também chamado de parâmetro de posição ou locação do item i , é medido na mesma escala da habilidade e representa o nível de qualidade de vida necessário para uma probabilidade de estar satisfeito igual a 0,5. Assim, quanto maior o valor de b_i mais difícil ou menor a probabilidade do idoso sentir-se satisfeito com o item em questão. À medida que b_i diminui, maior a probabilidade do idoso sentir-se satisfeito no item i . Valores de b_i negativos indicam que mais de 50% dos idosos com nível de qualidade de vida maior que b_i sentem-se satisfeitos com o item i , e valores de b_i positivos informam que menos de 50% dos idosos com qualidade de vida maior que b_i estão satisfeitos com o item analisado.

A correspondência entre o parâmetro de dificuldade e o traço latente consiste na premissa de que idosos com baixa qualidade de vida terão grande dificuldade (pouca chance) de estarem satisfeitos. Esta afirmação torna-se possível, pois ambos parâmetros estão na mesma escala, que representa uma das grandes vantagens da TRI sobre a teoria clássica. O parâmetro θ_j (habilidade) que representa a qualidade de vida latente do idoso j pode assumir qualquer valor, entretanto, é necessário estabelecer uma origem e uma unidade de medida para a definição da escala de qualidade de vida. Nestas análises optou-se pela escala com média 0 e desvio-padrão 1, a métrica mais simples e por consequência, mais empregada também.

Aspectos Éticos

Todos os pacientes incluídos no estudo de Moraes (2004) ¹ assinaram o Termo de Consentimento Informado, além de outros detalhes formais necessários.

Resultados

Pela Tabela 1 pode-se observar que a maioria dos respondentes é do sexo feminino (73,7%), e dentre as mulheres, grande parte possui de 60 a 69 anos (61%). Esta característica também aparece na descrição das faixas etárias para os homens (60%). Com relação à escolaridade, tanto para os homens (60%) quanto para as mulheres (76,5%) há predominância de nível educacional até o 1º grau completo. Grande parte das idosas é viúva (45,1%), ao passo que dentre os homens, a maioria (67,6%) são casados. Outra variável com semelhante distribuição de frequências entre homens e mulheres é a percepção sobre sua saúde estar boa ou nem ruim nem boa.

Escala de Qualidade de Vida de Flanagan (EQVF)

A EQVF mostrou-se fidedigna, com um valor para o α de Cronbach igual a 0,81, mostrando-se válida nesse contexto, tomada globalmente. Foi realizada uma análise fatorial exploratória (sem o item 5), indicando que a escala tem 3 fatores, o que contradiz o fato da elaboração original de 5 domínios/construtos. O primeiro fator foi composto pelos itens 9, 8, 13, 12, 10, 7, 11 e 14 e explica aproximadamente 30% da variação total. O segundo fator foi formado pelos itens 3, 6, 15 e 4 e o terceiro fator pelos itens 2 e 1, explicando, respectivamente, 9,4 e 7,6 % da variação total. Todos os fatores juntos explicaram aproximadamente 47% da variação total (Tabela 2 do Anexo H).

A EQVF, considerada multidimensional por natureza e construção, foi tomada globalmente, isto é, analisada sem levar em conta o domínio ao qual os itens pertenciam. Vale ressaltar que as informações foram dicotomizadas após a realização da pesquisa. Na Tabela 2 são fornecidas as estimativas para os parâmetros do ML-2, juntamente com seus respectivos erros padrões. Assim, dentre os 15 itens da escala, apenas 1 (item 5) foi retirado da mesma, não sendo considerado para as análises finais e cálculo das habilidades.

Assumindo que o nível de qualidade de vida pode ser representado pela distribuição normal padrão, valores negativos de dificuldade indicam que mais de 50% dos idosos com nível de qualidade de vida maior do que b_i sentem-se satisfeitos; já valores positivos para a dificuldade informam que menos de 50% dos indivíduos

estão satisfeitos. Isso indica o grau de dificuldade do idoso sentir-se satisfeito com determinada característica. Na Tabela 2, percebe-se que somente um item, relativo à aprendizagem, possuiu valor positivo para a dificuldade, o que caracteriza que menos de 50% dos idosos com nível de qualidade de vida superior a 0,556, estão satisfeitos com a sua aprendizagem. E ao analisar as estimativas para as discriminações, nota-se que os itens 10, 11, 12 e 13 são os itens com maior poder discriminativo, ou em outras palavras, são itens importantes para a relação de que pessoas com um bom nível de auto-conhecimento, emprego, boa comunicação criativa e oportunidades de lazer (recreação ativa) tendem a ter um nível superior de qualidade de vida. Na Figura 2 são apresentadas as curvas características dos itens (CCI) para todos os 15 itens, e fica nítido o motivo da retirada do item 5, já que sua curva não tem uma boa inclinação e comportamento totalmente diferente dos demais. O item que possui maior inclinação (curva mais íngreme) e por conseqüência, maior discriminação é o item 12.

Após a estimação através do ML-2 e geração de escores latentes referentes à qualidade de vida (Tabela 1 do Anexo G), sem o item 5, foi realizada uma análise de variância, para verificar quais fatores eram importantes no âmbito da qualidade de vida. Os fatores saúde percebida dicotomizada (1 = ruim, 2 = boa) e nível de depressão (1 = normal, 2 = leve a moderada, 3 = grave) mostraram-se significantes a 1%. Idosos com saúde percebida boa tendem a ter melhor qualidade de vida, assim como indivíduos não-deprimidos. Também foi realizada uma análise de variância, desta vez com os escores brutos, para os 14 itens, e assumindo-se um nível de significância de 1%, as mesmas variáveis mostraram-se significantes, com o acréscimo da variável relativa à desesperança, indicando que pessoas com grau de desesperança mínima, não-deprimidos e com saúde relatada boa tendem a ter um escore maior de qualidade de vida (Anexo I).

Foram computadas também as estatísticas clássicas para cada item da escala de qualidade de vida. Conforme a Tabela 3, pôde-se perceber que a maioria (85%) dos idosos sentiram-se satisfeitos com o item 4, relativo à construir família, assim como com ouvir música, assistir TV ou cinema (82%); e por outro lado, sentiram-se insatisfeitos com sua aprendizagem (65%). Esses achados concordam com os resultados da TRI, confirmando que os idosos tem maior dificuldade de se sentirem

satisfeitos com a sua aprendizagem, e menor dificuldade com o entretenimento e construir família. Os itens com maior correlação bisserial foram o 11 e 12, indicando que existe uma maior relação desses itens com o escore bruto, e que à medida que aumenta o escore bruto (com os 15 itens) aumentam também os escores para cada um desses itens. Esses mesmos itens são os que apresentaram maior poder de discriminação para a qualidade de vida, vindo à tona a relação que itens com correlação baixa (alta) apresentam baixa (alta) discriminação. O escore bruto global, com os 14 itens apresentou um escore médio de 9,7 com desvio-padrão de 3,3 unidades. Metade dos idosos possuem escores brutos maiores ou iguais a 10, e o escore mínimo é 0, e o máximo foi 14, representando, insatisfação total e satisfação total, respectivamente.

O escore bruto global apresentou uma correlação muito alta com o escore obtido via TRI, de magnitude 0,972 ($p < 0,0001$), indicando que à medida que aumentam os escores brutos, aumentam também, numa relação estritamente linear, os escores latentes.

Foi computada após a classificação pela mediana dos escores brutos e latentes, a estatística de concordância Kappa. Através dessa divisão pela mediana, apenas 13 idosos não foram classificados da mesma maneira. Esses indivíduos foram classificados como tendo baixa qualidade de vida pelo escore bruto, e alta qualidade de vida pelo escore latente. Os resultados indicam concordância entre os métodos de classificação ($kappa = 0,935$; Anexo J).

E ao serem comparados os resultados da TRI com os da Análise Fatorial (AF), percebeu-se que os itens se agruparam de acordo com a semelhança nos valores da discriminação, assim como acontece com as cargas da AF (Tabela 2 do Anexo H). Os itens com maiores cargas fatoriais no primeiro fator de uma maneira geral apresentaram os maiores valores estimados para o parâmetro de inclinação. Os resultados se assemelham fortemente.

WHOQOL-BREVE

O WHOQOL-BREVE mostrou-se ser uma escala fidedigna, seja ela tomada globalmente ou por domínios, com valores para o coeficiente α de Cronbach de 0,8011 e 0,8000, respectivamente. Foi realizada uma análise fatorial exploratória com os 23 itens consistentes, resultando numa estrutura com 6 fatores, resultado este

que não confirma o propósito do número de dimensões construídas, mesmo que os resultados não diverjam muito da estrutura elaborada pela OMS. O primeiro fator explicou aproximadamente 26% da variação, e os demais 5 fatores, explicaram entre 7 e 5%, somando 51,8% da variação total explicada pelos 6 fatores. No primeiro fator estão incluídos 10 itens, referentes à energia e atividades para o dia-a-dia, satisfação com saúde, tratamento médico, locomoção, capacidade pro trabalho, capacidade impedida pela dor, aparência física, vida sem sentido e lazer. Já no segundo fator estão inclusos a disponibilidade das informações e aproveitar a vida, no terceiro entraram a segurança, concentração e sentimentos negativos, e no quarto fator estão a satisfação com o local de moradia e transporte, ambiente físico e qualidade de vida em geral, e no último entrou o item de satisfação com o sono (Tabela 3 do Anexo H).

O modelo ML-2 foi ajustado à escala como um todo, tomada de forma global e por domínios. Na Tabela 4 são apresentadas as estimativas para os parâmetros de todos os itens. Ao ser considerada globalmente, 3 itens foram retirados da versão final por apresentarem pequenos valores de discriminação (itens F153, F181, F193), não mostrando-se apropriados para estimar qualidade de vida de idosos. Os itens retirados referem-se à vida sexual, dinheiro suficiente para satisfazer as necessidades e satisfação com acesso aos serviços de saúde. O primeiro item retirado pertence ao domínio social e os outros dois ao domínio meio-ambiente. Em suma, a versão final do instrumento global ficou com 23 itens. Os itens com poder de discriminação maior que 1 foram os itens F21, F91, F103, F124, F63 e G4, indicando que são apropriados para discriminar bem os indivíduos com maior ou menor qualidade de vida. O item F103 apresentou a maior discriminação (1,458), indicando que os idosos satisfeitos com suas capacidades de desempenharem suas atividades do dia-a-dia tiveram maiores escores de qualidade de vida. Na Figura 3 são apresentadas as CCI's para os 26 itens originais da escala, e pôde-se perceber que os 3 itens retirados possuem curvas pouco discriminativas a partir do ponto de inflexão. Esses itens foram retirados das análises das habilidades (traços latentes) finais. O item com inclinação mais acentuada foi o 8º item (F103), apresentando o maior poder de discriminação.

Todos os 23 itens remanescentes apresentaram dificuldades negativas, indicando que mais de 50% dos idosos com nível de qualidade de vida acima de b_i sentem-se satisfeitos com aqueles itens. O item que apresentou a maior dificuldade foi o item G4, que remete à satisfação de $-0,032$. Esse valor é o ponto na escala correspondente ao nível de qualidade de vida em que o idoso tem 0,5 de probabilidade de sentir-se satisfeito com a sua saúde em geral. O fato de todos os valores de dificuldades dos 23 itens restantes serem negativas indicaram que os idosos tem baixa dificuldade de estarem satisfeitos nos itens do instrumento.

Após o ajuste do ML-2 e geração de escores latentes (qualidade de vida), sem os três itens retirados, foi realizada uma análise de variância (Anexo I), para verificar quais fatores eram importantes no âmbito da qualidade de vida. Os fatores que se mostraram significantes foram a saúde percebida, níveis de depressão e desesperança. Os idosos que reportaram ter boa saúde tiveram um nível de qualidade de vida superior aos demais, assim como para os que foram diagnosticados sem depressão e grau de desesperança mínima. Também foi realizada uma análise de variância, desta vez com os escores brutos, para os 23 itens, e as mesmas variáveis mostraram-se significantes.

Foram calculadas também as estatísticas clássicas para os itens remanescentes. Conforme a Tabela 5, pôde-se perceber que a maioria dos idosos sentiram-se satisfeitos com o item F133, referente às relações pessoais, e sentiram-se mais insatisfeitos com o F153 (vida sexual). Essas afirmações confirmaram os resultados da TRI, onde os idosos apresentaram uma maior dificuldade de sentirem-se satisfeitos com sua vida sexual ($b = 0,905$), embora esse item tenha sido retirado das análises finais. E por outro lado, o valor baixo para a dificuldade do item F133, com $b = -1,916$, também concordou com os resultados clássicos.

Os itens com maior correlação bisserial foram os itens F103 e F124, que também apresentaram grandes valores para a discriminação. O escore bruto global com os 23 itens obteve uma média de 15,7 com um desvio-padrão de 5,2. Metade dos idosos teve escores brutos menores ou iguais a 16. O escore bruto mínimo foi 0 (insatisfação total) e o maior 23 (satisfação total).

Com o WHOQOL-BREVE, além de ser produzido um escore global com os 23 itens via TRI, também foi produzido um escore para cada domínio. O domínio 1

(físico), originalmente com 7 itens, apresentou 1 item inconsistente (F21), com poder de discriminação 0,391, inferior ao padrão estabelecido. E para os outros domínios (psicológico, social e meio ambiente) não houve retirada de nenhum item.

A fim de verificar a relação dos traços latentes gerados pelo modelo de TRI empregado, ora globalmente, ora por domínios calculou-se a correlação linear de Pearson. Todos os domínios se correlacionaram positiva e significativamente (a 1%) com o domínio global, e também entre si, apesar de algumas correlações serem de baixa magnitude, como mostra a Tabela 6. O escore de 23 itens apresentou uma maior correlação com o domínio físico ($r = 0,854$), seguido pelo domínio psicológico ($r = 0,787$), domínio social ($r = 0,488$) e domínio meio-ambiente ($r = 0,422$). O domínio físico correlacionou-se mais fortemente com o domínio psicológico, com um valor de correlação igual a 0,554, enquanto que o domínio psicológico manteve uma correlação da magnitude de 0,399 com o domínio social e de 0,361 com o meio ambiente. O domínio social apresentou correlação linear com o domínio ambiental ($r = 0,712$).

O escore bruto apresentou uma correlação com o escore obtido via TRI de magnitude 0,968 e $p < 0,0001$, indicando que à medida que aumentaram os escores brutos, aumentaram também os escores latentes.

Foi realizado, após a classificação pela mediana dos escores brutos e latentes, a estatística de concordância Kappa. Através dessa divisão pela mediana, 16 idosos não foram classificados da mesma maneira. Alguns indivíduos (9) foram classificados como tendo baixa qualidade de vida pelo escore bruto, e alta qualidade de vida pelo escore latente, e 7 idosos com o caso contrário. Verificou-se concordância da classificação dos idosos nos 2 grupos ($kappa = 0,92$).

E ao serem comparados os resultados da TRI com os da AF, percebeu-se que os itens se agruparam de acordo com a semelhança nas discriminações, assim como aconteceu com as cargas da AF. Os resultados foram muito semelhantes (Tabela 3 do Anexo H).

WHOQOL-100

A análise de fidedignidade (α de Cronbach) com os dados binários para o WHOQOL-100 resultou em um valor de 0,7711 levando em conta as facetas e um valor de 0,7613 para os domínios.

A análise fatorial exploratória para as facetas não confirmou a estrutura original de 6 domínios, e os resultados para as cargas fatoriais e estrutura diferem bastante dos resultados do Grupo WHOQOL. Isto era de se esperar, pois foram retirados 35 itens, o que alterou significativamente a estrutura originalmente elaborada. O primeiro fator explicou em torno de 33% da variação total e os demais de 7 a 5 % aproximadamente, somando 55,8% para todos os fatores (Tabela 4 do Anexo H).

Ao ser ajustado o ML-2 ao WHOQOL-100, outro instrumento multidimensional, tomado de forma global ou por domínios, ocorreram uma série de problemas, devido ao grande número de itens e parâmetros a serem estimados. O programa BILOG retirou automaticamente 19 itens (F11, F12, F13, F14, F22, F24, F81, F82, F83, F84, F93, F94, F102, F104, F131, F182, F184, F232, F234), pois apresentaram correlações bisseriais muito baixas. A esses 19 itens foram acrescentados mais 16 (F32, F34, F72, F73, F111, F112, F113, F114, F152, F153, F154, F163, F164, F181, F191, F222), com estimativas muito baixas para a discriminação ($a < 0,40$), resultando numa versão global final com 65 itens e 35 itens retirados. Na Figura 4 são apresentadas as curvas para os 81 itens iniciais, e dentre esses, os 16 itens que foram retirados devido às baixas inclinações.

A maioria dos itens restantes (Tabela 7) apresentou um bom poder discriminativo, com valores $a_i > 1$, sendo, portanto, bons medidores para discriminar/diferenciar os idosos com alta ou baixa qualidade de vida. Apenas os itens F51, F151 e F183, respectivamente, sobre a avaliação da memória, vida sexual e situação financeira, apresentaram estimativas para a dificuldade positiva e, todos os demais com índices negativos. O fato desses valores terem sido positivos indica que os idosos não estavam satisfeitos com a sua memória, vida sexual e situação financeira. Menos de 50% dos entrevistados com qualidade de vida acima de 1,210 sentiam-se satisfeitos na sua vida íntima. As interpretações são análogas para os demais.

Com os traços latentes computados e os escores brutos, sem os 35 itens inconsistentes, realizou-se uma análise de variância (Anexo I). Os fatores influentes foram os mesmos que os observados na análise do WHOQOL-BREVE: saúde percebida, nível de depressão e desesperança, tanto para os escores latentes, como para os escores brutos ($p < 0,0001$).

Foram calculadas também as estatísticas clássicas para os itens remanescentes. Conforme a Tabela 8, o item que os idosos mais se sentiam satisfeitos é o item F132, sobre os relacionamentos familiares e o item com que se sentiam mais insatisfeitos é o F73 sobre a sua aparência. O item F132 também apresentou o menor valor de dificuldade, concordando com os resultados das análises clássicas. Os itens com correlações bisseriais baixas foram os mesmos itens que já foram retirados da análise de TRI por apresentarem baixas discriminações. Novamente, os resultados da TRI e análise clássica concordaram em grande parte. A média para os escores brutos foi de 44,630 com mediana de 47, e máximo de 64. O escore mínimo encontrado foi 1.

O ML-2 também foi ajustado aos domínios, ou seja, foram ajustados 6 modelos, um para cada domínio, e calculados os escores latentes. O domínio físico produziu um escore latente com 8 itens. Já com o domínio psicológico, dos 20 itens, permaneceram 14 e do domínio 3 (nível de independência) restaram apenas 7 itens. No domínio social (12 itens) permaneceram 9 itens; no domínio 5 (ambiente) apenas 6 foram retirados; e do domínio referente à espiritualidade, nenhum item obteve estimativa inconsistente, e seu escore latente foi calculado considerando todos os 4 itens.

E, visando estimar a relação entre esses traços latentes gerados para cada um dos domínios e o escore latente global foi usada a análise de correlação linear de Pearson. O escore latente geral correlacionou-se significativamente com todos os domínios, mas de forma distinta: com os domínios físico, ambiente e espiritual de forma negativa, com correlações de -0,551, -0,297 e -0,159, respectivamente, e com os domínios psicológico ($r = 0,845$), nível de independência ($r = 0,725$) e domínio das relações sociais ($r = 0,497$) de forma positiva, como mostra a Tabela 9. O escore latente global também correlacionou-se com o escore bruto ($r = 0,940$).

E visando uma comparação do WHOQOL-BREVE com o WHOQOL-100, foi procedida uma análise de correlação linear com todos os escores latentes computados pela aplicação da TRI, tanto globalmente quanto por domínios. O escore latente do WHOQOL-BREVE apresentou uma correlação significativa com o escore produzido com o WHOQOL-100 ($r = 0,916$). Entre os domínios dos dois instrumentos obteve-se uma correlação de $-0,623$ no domínio físico, $0,785$ no psicológico, $0,750$ no social e de $-0,281$ no ambiental (Tabela 5 do Anexo H)

Foi calculada após a classificação pela mediana dos escores brutos e latentes, a estatística de concordância Kappa. Através dessa divisão pela mediana, 23 idosos não foram classificados da mesma maneira. Alguns indivíduos (12) foram classificados como tendo baixa qualidade de vida pelo escore bruto, e alta qualidade de vida pelo escore latente, e 11 com o caso contrário. Pôde-se concluir que pelos dois modos de escores os indivíduos são classificados de forma similar, apresentando concordância ($kappa = 0,885$; Anexo J).

E ao serem comparados os resultados da TRI com os da AF, percebeu -se que os itens se agruparam de acordo com a semelhança nos valores da discriminação, assim como acontece com as cargas da AF. No primeiro fator da AF estão os itens mais discriminantes na TRI. Os resultados foram muito semelhantes (Tabela 4 do Anexo H).

Na Figura 5, encontram-se os histogramas com a aderência à normalidade das habilidades (traços latentes) e dos escores brutos. Os escores latentes apresentaram uma maior aderência à curva normal, visto que os mesmos foram computados para ter média 0 e desvio-padrão 1. Por outro lado, os escores brutos apresentaram comportamento assimétrico. As maiores violações de normalidade aconteceram na EQVF, provavelmente devido ao pequeno número de itens presentes na escala. Fato contrário aconteceu com o WHOQOL-100, cujos escores brutos apresentaram uma aderência mais acentuada.

Discussão

Neste trabalho, analisaram-se os resultados da avaliação da qualidade de vida de idosos socialmente ativos obtidos através dos instrumentos WHOQOL-100, WHOQOL-BREVE e EQVF, usando a metodologia de TRI, bem como compararam-se os resultados obtidos com os gerados pelos métodos tradicionais de análise, a TCM e a análise fatorial (AF).

O escore total, via TCM, ou escores fatoriais, na AF, podem ser representados pelo escore latente, obtido na TRI. A grande desvantagem do escore total ou dos escores produzidos para cada um dos fatores da AF, é que esses escores levam em conta todos os itens, não diferenciando quais são mais ou menos relevantes para a estimação da qualidade de vida, como acontece na TRI. Tal importância/relevância do item é medida pelo parâmetro da discriminação^{34,35}, que substitui os tão usados α de Cronbach e correlações bisseriais. A TRI separa os parâmetros estimados para o item dos estimados para os indivíduos.

Os instrumentos utilizados, WHOQOL-100, WHOQOL-BREVE e EQVF^{11,19-21,23,24} são multidimensionais por natureza e construção, não devendo ser analisados unicamente de forma global, mas também por domínios. Todos foram validados para idosos^{1,20}. Mesmo com essas evidências em mente, a EQVF foi avaliada unicamente de forma global, pois não seria possível, nem correto, aplicar o modelo de TRI a domínios com apenas 2 itens. A escala também foi dicotomizada, visando resolver problemas de ajuste aos modelos de TRI, muito embora a decorrente perda de informações. Os resultados evidenciaram a necessidade da retirada do item 5 (relacionamento íntimo), resultado este que não confirma os obtidos por outros autores, que usaram outras técnicas. Não há como serem feitas comparações, devido ao pioneirismo de aplicar TRI na EQVF. O item referente ao relacionamento íntimo merece uma atenção especial, e para uma melhor análise desta informação recomenda-se a incorporação de mais um parâmetro no modelo, conhecido como parâmetro para itens constrangedores³⁴.

Existe a necessidade de adaptação da EQVF para idosos ativos, visto que a mesma é amplamente usada em pacientes com doenças crônicas^{1,23,24}. O fato dos idosos, principalmente os ativos, superestimarem a sua qualidade de vida poderia

comprometer grande parte dos resultados encontrados. Moraes (2004) utilizando a escala com 7 pontos também não confirmou a estrutura original da escala, com 5 fatores, e seus resultados concordam com os achados neste trabalho¹. Outros autores usando AF, excluíram os itens 7 (voluntariado) e 14 (entretenimento) devido às baixas cargas fatoriais. Seus resultados pouco concordam com a forma com que a escala foi planejada²⁴.

A relação entre os parâmetros de inclinação/discriminação com as correlações bisseriais e as cargas fatoriais foi verificada na prática. Itens com correlações bisseriais baixas resultaram em estimativas pobres para a discriminação, ao passo que itens agrupados em fatores apresentaram valores semelhantes para as discriminações. Ficou claro também ao serem comparadas as porcentagens de satisfação/insatisfação com as estimativas para as dificuldades, ressaltando a idéia de que itens com baixos percentuais de satisfação resultam em estimativas para a dificuldade maiores e vice-versa.

Os resultados obtidos nas análises clássicas e fatorial para o WHOQOL-BREVE são muito similares aos resultados da TRI, inclusive na classificação das habilidades, não havendo diferenças importantes na categorização de qualidade de vida. O escore latente global se correlacionou de forma mais intensa com os escores produzidos pelos domínios físico e psicológico. Esses resultados concordam em parte com os achados por Pereira et al. (2006)²⁰, ainda que os autores tenham usado outras técnicas, ao invés do escore via TRI. Segundo os autores, o domínio que mais contribuiu para a qualidade de vida global foi o físico, seguido do ambiental e do psicológico, evidenciando o fato de que os domínios explicam limitadamente a qualidade de vida global, mas alterações em um ou mais domínios podem implicar em alterações na qualidade de vida global²⁰. O domínio referente à capacidade funcional assume grande relevância no conceito multidimensional de qualidade de vida, devendo ser medido juntamente com o estado de saúde quando se deseja avaliar qualidade de vida de idosos. Dessa forma, o bem-estar do idoso seria resultado do equilíbrio entre as diversas dimensões, tendo em vista que o bem-estar pode ser atingido por muitos, independentemente da presença ou não de enfermidades.

A maior influência do domínio físico na qualidade de vida global dos idosos, verificada no presente trabalho, medida pelas correlações lineares com o escore

latente, ressalta a importância de se considerar a capacidade funcional como importante fator de impacto na qualidade de vida em idosos. A capacidade funcional, atualmente, surge como um novo paradigma de saúde para os indivíduos idosos, e o envelhecimento saudável passa a ser visto como uma interação multidimensional entre saúde física e mental, independência na vida diária, integração social, suporte familiar e independência econômica. Além disso, a qualidade de vida na velhice tem sido associada a questões de independência e autonomia, e a dependência do idoso resulta das alterações biológicas (incapacidades) e de mudanças nas exigências sociais²⁰.

O WHOQOL-100 perdeu sua configuração ao serem retirados 35 itens inconsistentes identificados pelos métodos de TRI, mostrando-se inadequado globalmente, exigindo grande cautela nas interpretações e conclusões. As análises por domínios também não foram de todo satisfatórias, onde novamente ocorreu acentuada retirada de itens inconsistentes. Há evidências de que essa estrutura não é a mais adequada para estimar qualidade de vida em idosos, ressaltando a necessidade de serem criados instrumentos próprios para idosos, especialmente os ativos. Um exemplo de instrumento com perspectiva transcultural e específico para idosos é o WHOQOL-OLD, com boas características psicométricas e validado para a população brasileira⁴.

O escore latente global via TRI se correlacionou de forma mais intensa com os domínios psicológico e nível de independência, indicando a importância que o idoso ativo dá ao fato de não depender dos outros. Contrariando a expectativa dos resultados, o escore latente apresentou correlação negativa com o domínio físico, além dos domínios social, espiritual e com a faceta geral. Esses resultados podem ser devidos ao grande número de itens inconsistentes, seja de forma global ou por domínios. Os domínios físicos do WHOQOL-100 e WHOQOL-BREVE se correlacionaram de forma negativa, dando evidências de que não estão medindo a qualidade de vida da mesma forma. Esses resultados podem ser devidos em parte à grande retirada de itens inconsistentes, dentro de cada domínio do WHOQOL-100, contrariando os resultados esperados. Os domínios psicológico e social para as duas escalas se correlacionaram positivamente. Neste estudo o WHOQOL-BREVE não mostrou ser uma opção altamente satisfatória frente ao instrumento mais longo.

Maior atenção e cautela deve ser dada a essas interpretações de equivalência entre os instrumentos. Os resultados não concordam com os achados por Noerholm et al. (2004)²¹ e O'Carroll et al. (2000)²², que afirmam que cada domínio do WHOQOL-BREVE é uma escala válida e fidedigna para medir os respectivos domínios do WHOQOL-100, e que cada domínio é uma escala unidimensional.

A interpretação da avaliação global da qualidade de vida é um tanto difícil, uma vez que resultados semelhantes podem ser explicados por diferentes fatores. Para cada indivíduo, há uma forma de operacionalizar sua avaliação, e a avaliação de um mesmo indivíduo pode variar com o tempo, com a variação de prioridades ao longo da vida e com as circunstâncias pelas quais a vida pode se modificar. Assim, a qualidade de vida global e os domínios da qualidade de vida parecem ser duas maneiras diferentes de se avaliar a qualidade de vida de idosos. Isso é especialmente importante no delineamento de pesquisas, uma vez que a medida a ser utilizada irá depender dos objetivos da pesquisa para avaliar a qualidade de vida com um todo (domínio global) ou apenas alguns aspectos da qualidade de vida.

Outra questão pertinente na avaliação da qualidade de vida é a hipótese que os idosos tendem a baixar as suas expectativas e, em geral, aceitam como fato normal as suas limitações físicas ou materiais. O grau de conformismo pode, quem sabe, constituir um indicador de defesa ou de fuga das decepções. O idoso é um produto do meio em que vive, e interage com os fatores individuais, familiares, sócio-culturais e espirituais¹.

O estudo apresenta algumas limitações próprias dos instrumentos utilizados, como serem multidimensionais por natureza e construção, levando à escolha de modelos multidimensionais, e a ocorrência de itens com dois pólos no WHOQOL-100 e WHOQOL-BREVE, que devem ter seus escores transformados, mas que na aplicação do questionário podem gerar confusão, visto que fazem com que o raciocínio do respondente mude de direção várias vezes. Esses itens com dois pólos medem a qualidade de vida, mas em direções contrárias e também merecem atenção e cautela nas suas interpretações. Outra questão pertinente é a construção do instrumento abreviado a partir das respostas ao WHOQOL-100; este último aplicado aos idosos; podendo produzir assim respostas diferentes caso fosse aplicado diretamente.

A dicotomização das respostas, que por um lado resultou na perda de informações, mas só assim viabilizou a aplicação do modelo de TRI, pode ser uma limitação própria do estudo, além da grande retirada dos itens para comparar diretamente os resultados entre as técnicas. Outro aspecto relevante é a grande porcentagem de mulheres presentes no estudo (73,8%), podendo o sexo ser um fator de confundimento, embora não tenha sido significativo em nenhuma análise, tampouco na verificação de comportamento diferencial (DIF) entre os itens.

O modelo de TRI usado neste trabalho mostrou ser uma ferramenta útil e promissora na avaliação da qualidade de vida, como alternativa ou tomada de forma conjunta com a TCM e AF. As interpretações dos parâmetros foram satisfatórias e análogas às outras análises. Futuras pesquisas são necessárias para disseminar o uso da TRI em análises de instrumentos de avaliação, especialmente, os de qualidade de vida, e conseqüente interpretação das escalas de qualidade de vida obtidas via TRI.

Colaboradores

K. Guewehr foi responsável pelas análises, interpretação e redação do artigo e J. Riboldi colaborou na obtenção dos dados com o responsável, orientação, redação e revisão do artigo.

Referências

1. Moraes JF. Fatores determinantes do envelhecimento bem-sucedido do idoso socialmente ativo da região metropolitana de Porto Alegre [Tese de Doutorado]. Porto Alegre (RS). Programa de Pós-graduação em Gerontologia Médica, Pontifícia Universidade Católica do Rio Grande do Sul; 2004.
2. Fundação Instituto Brasileiro de Geografia e Estatística. Download: Estatística: perfil dos idosos responsáveis pelo domicílio. <http://www.ibge.gov.br> (acessado em 1º / dez/ 2006)
3. Veras RP, Ramos LR, Kalache A. Crescimento da população idosa no Brasil: transformações e conseqüências na sociedade. *Revista de Saúde Pública* 1987; 21 (3): 225-233.
4. Fleck MPA, Chachamovich E, Trentini C. Development and validation of the Portuguese version of the WHOQOL-OLD module. *Revista de Saúde Pública* 2006; 40(5): 785-791.
5. Ramos LR, Veras RP, Kalache A. Envelhecimento populacional: uma realidade brasileira. *Revista de Saúde Pública* 1987; 21(3): 211-224.
6. Kalache A, Veras RP, Ramos LR. O envelhecimento da população mundial: um desafio novo. *Revista de Saúde Pública* 1987; 21(3): 200-210.
7. Chaimovicz F. A saúde dos idosos brasileiros às vésperas do século XXI: problemas, projeções e alternativas. *Revista de Saúde Pública* 1997; 31(2): 184-200.
8. Paschoal SMP. Qualidade de vida do idoso: elaboração de um instrumento que privilegia sua opinião [Dissertação de Mestrado]. São Paulo (SP). Faculdade de Medicina, Universidade de São Paulo; 2000.
9. Ramos LR. Fatores determinantes no envelhecimento saudável em idosos residentes em centro urbano: Projeto Epidoso, São Paulo. *Cadernos de Saúde Pública* 2003; 19(3): 793-798.
10. Sousa L, Galante H, Figueiredo D. Qualidade de vida e bem-estar de idosos: um estudo exploratório na população portuguesa. *Revista de Saúde Pública* 2003; 37: 364-371.
11. Fleck MPA, Leal OF, Louzada S., Xavier M., Chachamovich E., Vieira G. et al. Desenvolvimento da versão em português do instrumento de avaliação de qualidade de vida da OMS (WHOQOL – 100). *Revista Brasileira de Psiquiatria* 1999; 21(1): 19-28.

12. Vecchia RD, Ruiz T, Bocchi SCM, Corrente JE. Qualidade de vida na terceira idade: um conceito subjetivo. *Revista Brasileira de Epidemiologia* 2005; 8(3): 246-252.
13. Farquhar M. Elderly people's definitions of quality of life. *Social Science and Medicine* 1995; 41(10): 1439-1446.
14. Bowling A, Banister D, Sutton S, Evans O, Windsor J. A multidimensional model of the quality of life in older age. *Aging and Mental Health* 2002; 6(4): 355-371.
15. Browne JP, O'Boyle CA, McGee HM, Joyce CR, McDonald NJ, O'Malley K, et al. Individual quality of life in the health elderly. *Quality of Life Research* 1994; 3(4): 235-244.
16. Xavier FMF, Ferraz MPT, Marc N, Escosteguy NU, Moriguchi EH. Elderly people's definition of quality of life. *Revista Brasileira de Psiquiatria* 2003; 25(1): 31-39.
17. Berlim MT, Fleck MPA. "Quality of Life": a brand new concept for research and practice in psychiatry. *Revista Brasileira de Psiquiatria* 2003; 25(4): 249-252.
18. Fleck MPA, Borges ZN, Bolognesi G, Rocha NS. Desenvolvimento do WHOQOL módulo espiritualidade, religiosidade e crenças pessoais. *Revista de Saúde Pública* 2003; 37(4): 446-455.
19. Fleck MPA, Leal OF, Louzada S., Xavier M., Chachamovich E., Vieira G., Santos L, et al. Aplicação da versão em português do instrumento abreviado de avaliação da qualidade de vida "WHOQOL – Bref". *Revista de Saúde Pública* 2000; 34(2): 178-183.
20. Pereira RJ, Cotta RMM, Franceschini SCC, Ribeiro RCL, Sampaio RF, Priore SE, Cecon PR. Contribuição dos domínios físico, social, psicológico e ambiental para a qualidade de vida global de idosos. *Revista de Psiquiatria do Rio Grande do Sul* 2006; 28(1): 27-38.
21. Noerholm V, Groenvold M, Watt T, Bjorner JB, Rasmussen NA, Bech P. Quality of life in the Danish general population-normative data and validity of WHOQOL-Bref using Rasch and item response theory models. *Quality of Life Research* 2004; 13: 531-540.
22. O'Carroll, Smith K, Coustan M, Cossar JA, Hayes PC. A comparison of the WHOQOL-BREF in detecting change in quality of life following liver transplantation. *Quality of Life Research* 2000; 9:121-124.
23. Lentz RS, Costenaso RGS, Gonçalves LHT, Nassar SM. O profissional de enfermagem e a qualidade de vida: uma abordagem fundamentada nas

- dimensões propostas por Flanagan. *Revista Latino-americana de Enfermagem* 2000; 8(4): 7-14.
24. Santos SR, Santos IBC, Fernandes MGM, Henrique MERM. Qualidade de vida do idoso na comunidade: aplicação da Escala de Flanagan. *Revista Latino-americana de Enfermagem* 2002; 10(6): 757-764.
 25. Power M, Quinn K, Schmidt S; WHOQOL-OLD Group. Development of the WHOQOL-Old module. *Quality of Life Research* 2005; 14(10): 2197-2214.
 26. Fleck MPA, Chachamovich E, Trentini CM. Projeto WHOQOL-OLD: método e resultado de grupos focais no Brasil. *Revista de Saúde Pública* 2003; 37(6): 793-799.
 27. Vianna HM. *Testes em Educação*. São Paulo: IBRASA – Instituição Brasileira de Difusão Cultural S.A.; 1978.
 28. Pasquali L. *Psicometria: Teoria e Aplicações*. Brasília: Editora Universidade de Brasília, 1997.
 29. Andrade DF, Tavares HR, Valle RC. *Teoria de Resposta do Item: Conceitos e Aplicações*. São Paulo: Associação Brasileira de Estatística; 2000.
 30. Presa DS. *A teoria de resposta ao item na avaliação educacional: Uma análise das questões que compõem a prova de matemática do concurso vestibular 2004 – UFRGS [Monografia]*. Porto Alegre (RS). Departamento de Estatística, Universidade Federal do Rio Grande do Sul; 2005.
 31. Valle RC. *Teoria de Resposta ao Item [Dissertação de Mestrado]*. São Paulo (SP). Instituto de Matemática e Estatística, Universidade de São Paulo; 1999.
 32. Embretson SE, Reise. *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum Associates; 2000.
 33. Mesbach M, Cole BF, Lee M-LT. *Statistical Methods for Quality of Life Studies*. London: Kluwer Academic Publishers; 2002.
 34. Cúri M. *Análise de questionários com itens constrangedores [Tese de Doutorado]*. São Paulo (SP). Instituto de Matemática e Estatística, Universidade de São Paulo; 2006.
 35. Vergara LGL. *Avaliação do Ensino de Ergonomia para o Design aplicando a Teoria de Resposta ao Item (TRI) [Tese de Doutorado]*. Florianópolis (SC) Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina; 2005.
 36. Alexandre JWC, Andrade DF, Vasconcelos AP, Araújo AMS. *Uma proposta de análise de um construto para medição dos fatores críticos da gestão pela*

qualidade por intermédio da Teoria de Resposta ao Item. *Gestão e Produção* 2002; 9(2): 129-141.

37. Soares TM. Utilização da Teoria de Resposta ao Item na produção de indicadores sócio-econômicos. *Pesquisa Operacional* 2005; 25(1): 83-112.
38. Cunha JA. Manual da versão em português das escalas Beck. São Paulo: Casa do Psicólogo; 2001.
39. Mislevy RJ, Bock RD. *Bilog 3: Item Analysis and Test Scoring with Binary Logistic Models*. Chicago: Scientific Software, Inc; 1990.

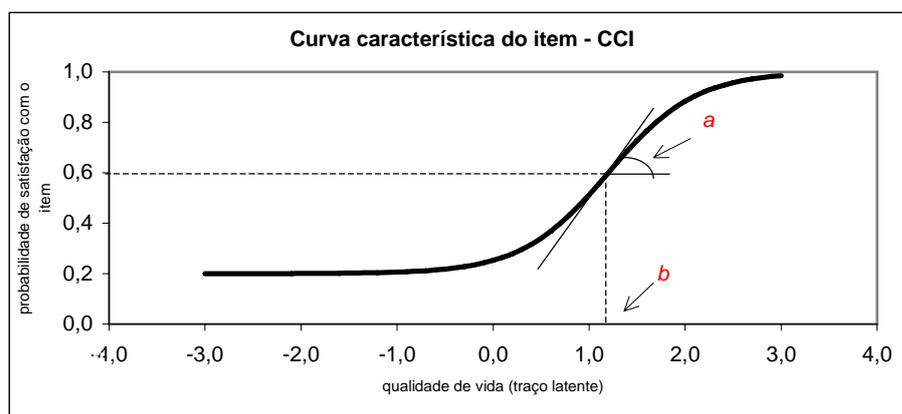


FIGURA 1 - Curva Característica do Item para Modelo Logístico de 2 Parâmetros

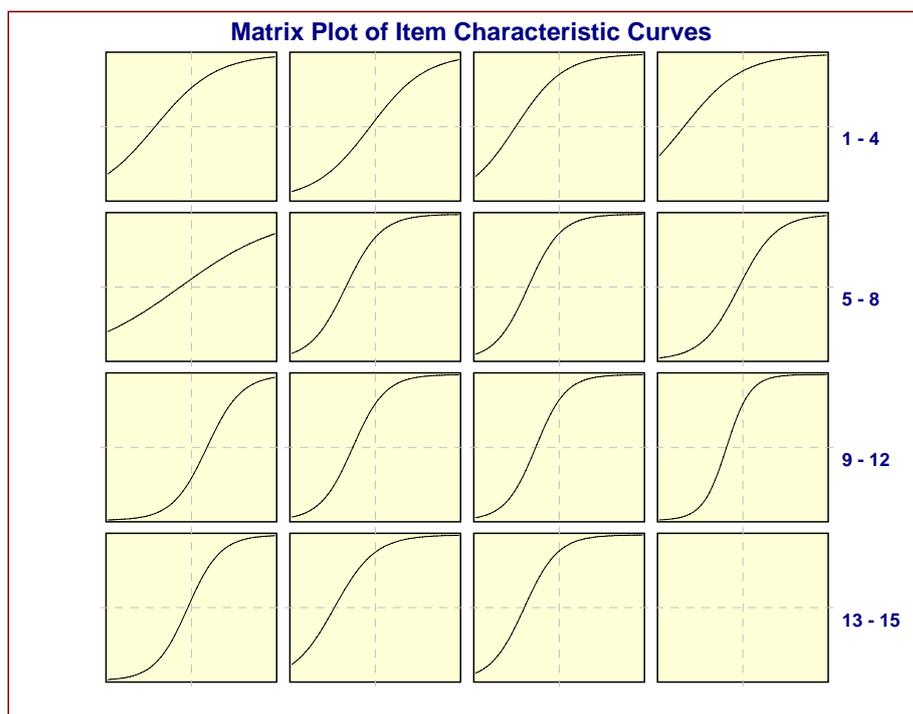


FIGURA 2 – Curvas características dos itens da EQVF

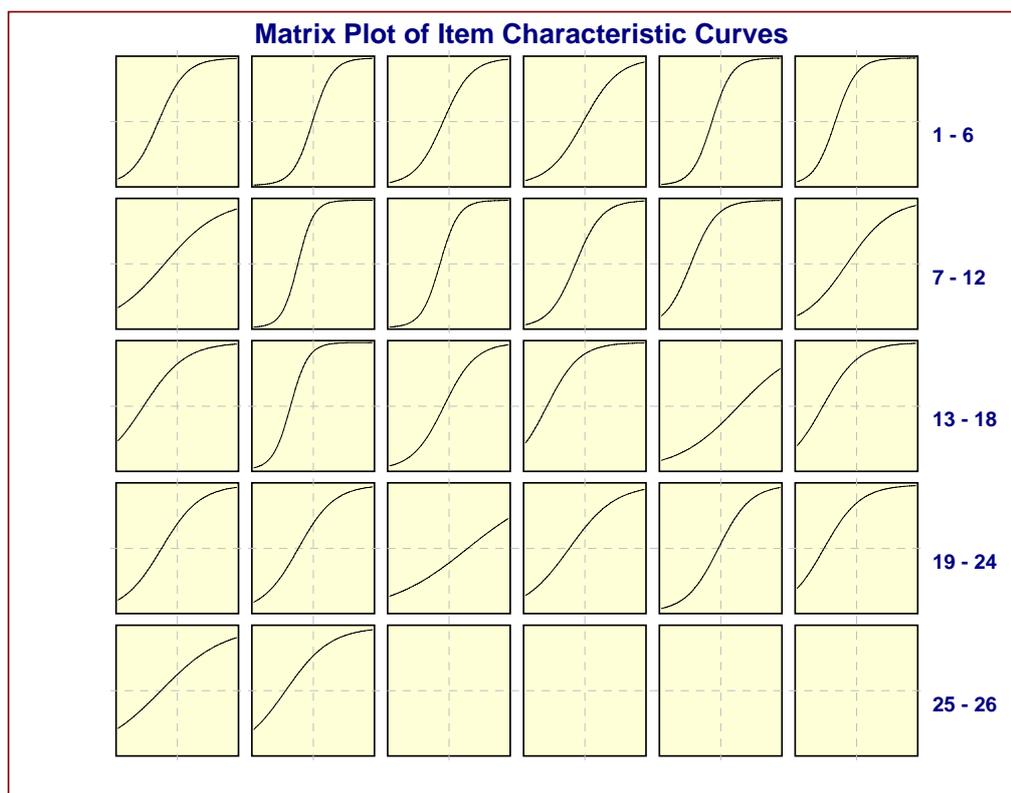


FIGURA 3 – Curvas características dos itens do WHOQOL-BREVE (G1, G4, F14, F113, F21, F91, F33, F103, F124, F41, F242, F53, F71, F63, F81, F133, F153, F144, F161, F221, F181, F201, F211, F173, F193, F233)

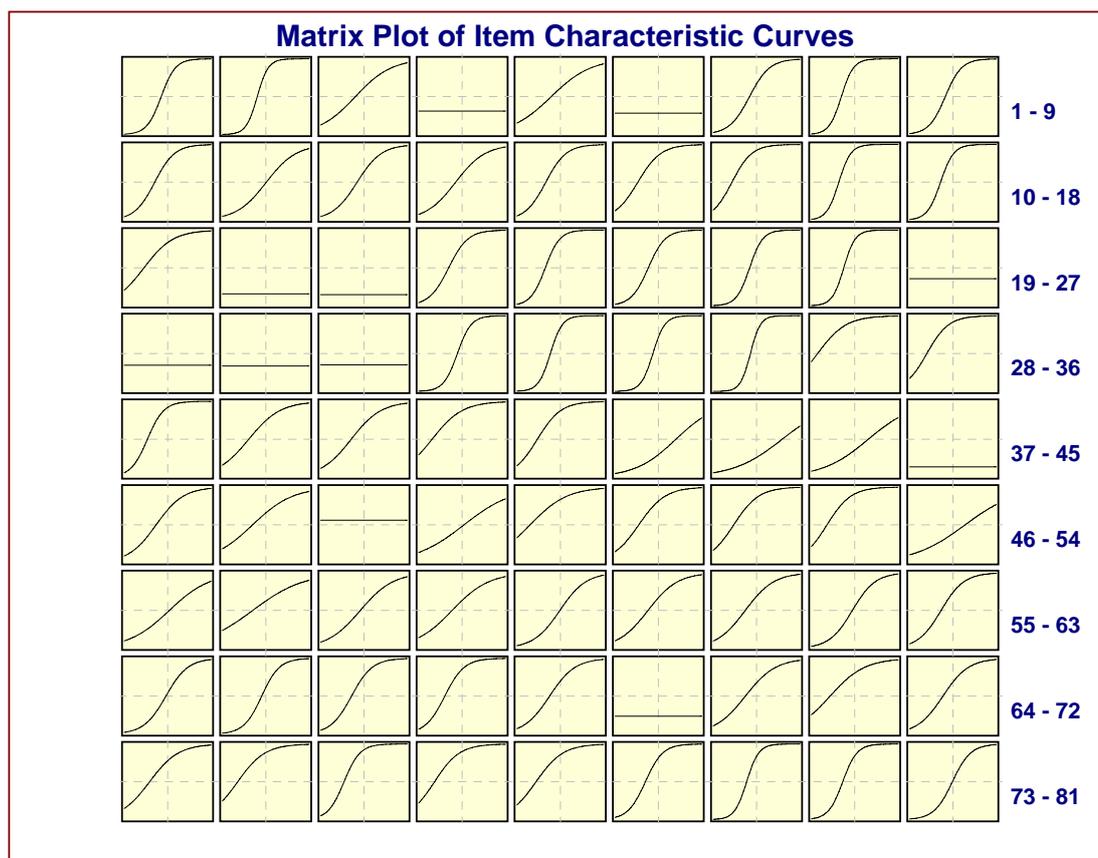


FIGURA 4 – Curvas características dos itens do WHOQOL-100 (F21, F23, F31, F32, F33, F34, F41, F42, F43, F44, F51, F52, F53, F54, F61, F62, F63, F64, F71, F72, F73, F74, F91, F92, F101, F103, F111, F112, F113, F114, F121, F122, F123, F124, F132, F133, F134, F141, F142, F143, F144, F151, F152, F153, F154, F161, F162, F163, F164, F171, F172, F173, F174, F181, F183, F191, F192, F193, F194, F201, F202, F203, F204, F211, F212, F213, F214, F221, F222, F223, F224, F231, F233, F241, F242, F243, F244, F251, F252, F253, F254)

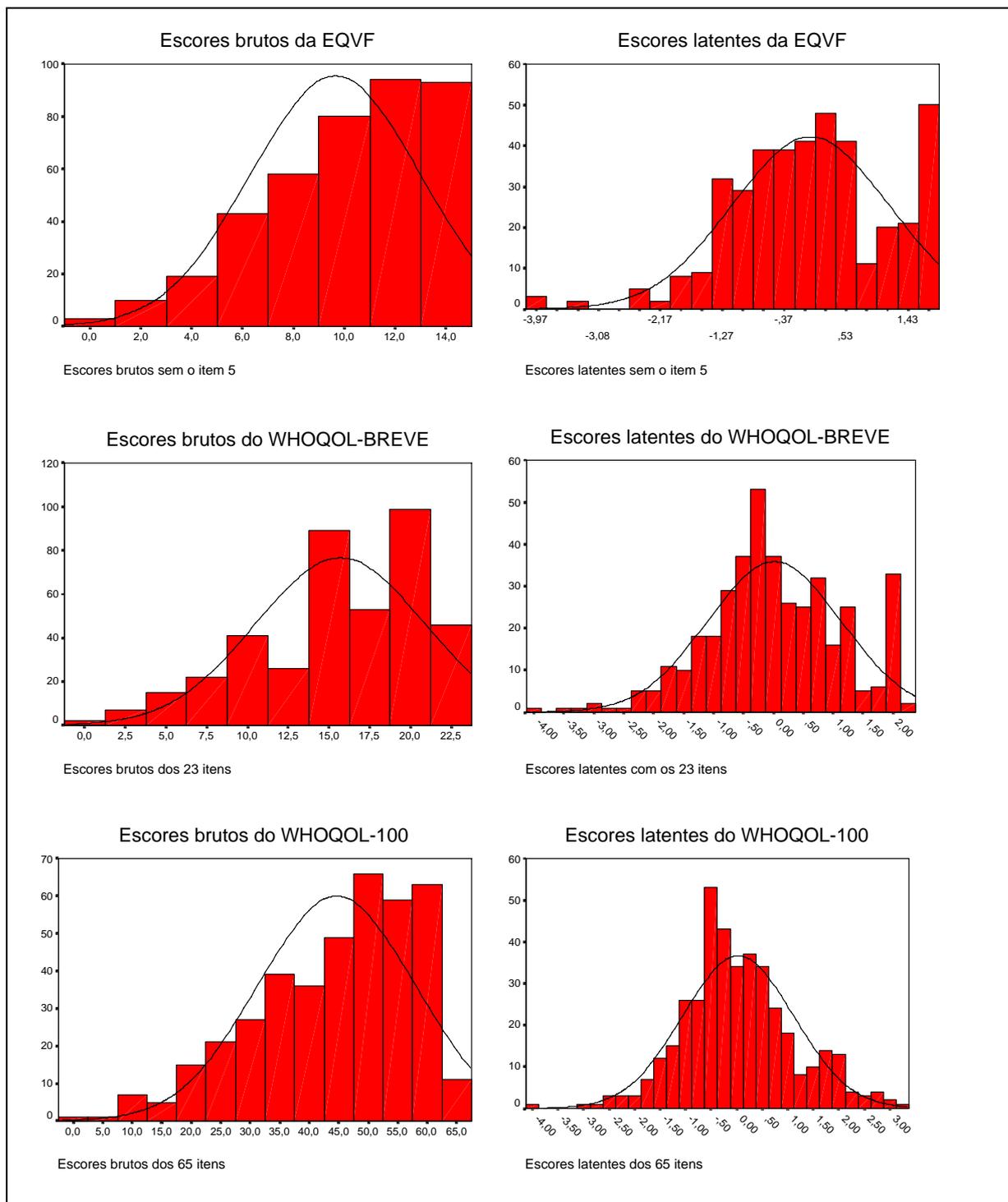


FIGURA 5 – Distribuição dos escores brutos e latentes para a EQVF, WHOQOL-BREVE e WHOQOL-100

TABELA 1 - Descrição dos dados dos 400 idosos socialmente ativos da região metropolitana de Porto Alegre

Sexo	Total	Feminino	Masculino
	%	73,7	26,3
Faixa etária (em anos)	60-64	93 (31,5%)	35 (33,3%)
	65-69	87 (29,5%)	28 (26,7%)
	70-74	64 (21,7%)	24 (22,9%)
	75-79	35 (11,9%)	13 (12,4%)
	80 e mais	16 (5,4%)	5 (4,8%)
Escolaridade	Analfabeto	32 (10,8%)	6 (5,7%)
	1º grau incompleto	144 (48,8%)	36 (34,3%)
	1º grau completo	50 (16,9%)	21 (20,0%)
	2º grau incompleto	0 (0%)	0 (0%)
	2º grau completo	48 (16,3%)	27 (25,7%)
	3º grau incompleto	4 (1,4%)	2 (1,9%)
	3º grau completo	13 (4,4%)	7 (6,7%)
	Pós-Graduação incompleta	0 (0%)	1 (1,0%)
	Pós-Graduação completa	4 (1,4%)	5 (4,8%)
Estado civil	Solteiro	11 (3,7%)	2 (1,9%)
	Casado	114 (38,6%)	71 (67,6%)
	Vivendo como casado	2 (0,7%)	2 (1,9%)
	Separado	15 (5,1%)	6 (5,7%)
	Divorciado	20 (6,8%)	8 (7,6%)
	Viúvo	133 (45,1%)	16 (15,2%)
Como está sua saúde	Muito ruim	21 (7,1%)	2 (1,9%)
	Fraca	34 (11,5%)	15 (14,3%)
	Nem ruim nem boa	103 (34,9%)	35 (33,3%)
	Boa	107 (36,3%)	40 (38,1%)
	Muito boa	30 (10,2%)	13 (12,4%)

TABELA 2 – Distribuição dos itens da Escala de Qualidade de Vida de Flanagan, estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses

Elementos da Escala de Qualidade de Vida de Flanagan (itens)	Estimativas dos parâmetros	
	a (discriminação)	b (dificuldade)
Bem estar físico e mental		
1. Conforto material: casa, alimentação, situação financeira	0,535 (0,103)	-1,314 (0,244)
2. Saúde: fisicamente bem e vigoroso(a)	0,590 (0,097)	-0,186 (0,130)
Relações com outras pessoas		
3. Relacionamento com pais, irmãos e outros parentes: comunicação, visita e ajuda	0,680 (0,118)	-1,565 (0,247)
4. Construir família: ter e criar filhos	0,561 (0,126)	-2,125 (0,427)
5. Relacionamento íntimo com esposo(a), namorado(a) ou outra pessoas relevante	0,322 (0,079)	-0,415 (0,232)
6. Amigos próximos: compartilhar interesses, atividades e opiniões	0,924 (0,156)	-1,063 (0,153)
Atividades sociais, comunitárias e cívicas		
7. Voluntariamente, ajudar e apoiar as pessoas	0,993 (0,180)	-1,119 (0,146)
8. Participação em associações e atividades de interesse público	0,874 (0,137)	-0,150 (0,102)
Desenvolvimento pessoal e realização		
9. Aprendizagem: freqüentar outros cursos para conhecimentos gerais	0,949 (0,144)	0,556 (0,109)
10. Auto-conhecimento: reconhecer seus potenciais e limitações	1,009 (0,158)	-0,810 (0,120)
11. Trabalho (emprego ou em casa): atividade interessante, gratificante e que vale a pena	1,087 (0,161)	-0,845 (0,112)
12. Comunicação criativa	1,385 (0,208)	-0,592 (0,091)
Recreação		
13. Participação em recreação ativa	1,004 (0,150)	-0,134 (0,092)
14. Ouvir música, assistir TV ou cinema, leitura ou outros entretenimentos	0,801 (0,137)	-1,474 (0,199)
15. Socialização: “fazer amigos”	0,979 (0,175)	-1,254 (0,164)

TABELA 3 – Estatísticas clássicas para os itens da EQVF

Item	Número de idosos satisfeitos (%)	Número de idosos insatisfeitos (%)	Correlação de Pearson	Correlação Bisserial
1	294 (73,5)	106 (26,5)	0,333	0,448
2	215 (53,8)	185 (46,2)	0,376	0,472
3	324 (81,0)	76 (19,0)	0,380	0,549
4	340 (85,0)	60 (15,0)	0,313	0,480
5	221 (55,2)	179 (44,8)	0,250	0,315
6	305 (76,2)	95 (23,8)	0,481	0,662
7	313 (78,2)	87 (21,8)	0,462	0,647
8	215 (53,8)	185 (46,2)	0,448	0,563
9	140 (35,0)	260 (65,0)	0,418	0,539
10	286 (71,5)	114 (28,5)	0,484	0,644
11	292 (73,0)	108 (27,0)	0,504	0,677
12	272 (68,0)	128 (32,0)	0,568	0,740
13	214 (53,5)	186 (46,5)	0,494	0,619
14	328 (82,0)	72 (18,0)	0,395	0,579
15	323 (80,8)	77 (19,2)	0,440	0,634

Estatísticas descritivas dos escores brutos da EQVF		
	Global (15 itens)	Global (sem item 5)
Média	10,205	9,652
Desvio-padrão	3,493	3,335
Mínimo	0	0
Máximo	15,0	14,0
Mediana	11,0	10,0

TABELA 4 - Distribuição dos itens do Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS – versão abreviada (WHOQOL-BREVE), estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses

Elementos do WHOQOL-BREVE (itens)	Estimativas dos parâmetros	
	a (discriminação)	b (dificuldade)
Domínio físico		
F14. Em que medida você acha que sua dor (física) impede você de fazer o que você precisa?	0,824 (0,107)	-0,299 (0,087)
F113. O quanto você precisa de algum tratamento médico para levar sua vida diária?	0,655 (0,094)	-0,081 (0,103)
F21. Você tem energia suficiente para seu dia-a-dia?	1,168 (0,143)	-0,443 (0,070)
F91. Quão bem você é capaz de se locomover?	1,086 (0,149)	-1,080 (0,109)
F33. Quão satisfeito(a) você está com o seu sono?	0,421 (0,079)	-0,649 (0,176)
F103. Quão satisfeito(a) você está com sua capacidade de desempenhar as atividades do seu dia-a-dia?	1,458 (0,223)	-0,776 (0,073)
F124. Quão satisfeito(a) você está com sua capacidade para o trabalho?	1,249 (0,152)	-0,458 (0,069)
Domínio psicológico		
F41. O quanto você aproveita a vida?	0,878 (0,125)	-0,475 (0,092)
F242. Em que medida você acha que a sua vida tem sentido?	0,906 (0,148)	-1,510 (0,171)
F53. O quanto você consegue se concentrar?	0,531 (0,084)	-0,489 (0,136)
F71. Você é capaz de aceitar sua aparência física?	0,554 (0,101)	-1,707 (0,262)
F63. Quão satisfeito(a) você está consigo mesmo?	1,314 (0,196)	-1,147 (0,099)
F81. Com que frequência você tem sentimentos negativos tais como mau humor, desespero, ansiedade, depressão?	0,735 (0,104)	-0,255 (0,097)
Domínio social (relações sociais)		
F133. Quão satisfeito(a) você está com suas relações pessoais (amigos, parentes, conhecidos, colegas)?	0,715 (0,132)	-1,916 (0,267)
F153. Quão satisfeito(a) você está com sua vida sexual?	0,379 (0,076)	0,905 (0,231)
F144. Quão satisfeito(a) você está com o apoio que você recebe de seus amigos?	0,651 (0,112)	-1,698 (0,240)

TABELA 4 - Distribuição dos itens do Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS – versão abreviada (WHOQOL-BREVE), estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses (continuação)

Elementos do WHOQOL-BREVE (itens)	Estimativas dos parâmetros	
	a (discriminação)	b (dificuldade)
Domínio meio-ambiente		
F161. Quão seguro(a) você se sente em sua vida diária?	0,603 (0,094)	-0,801 (0,145)
F221. Quão saudável é o seu ambiente físico (clima, barulho, poluição, atrativos)?	0,645 (0,099)	-0,753 (0,133)
F181. Você tem dinheiro suficiente para satisfazer suas necessidades?	0,291 (0,070)	0,936 (0,301)
F201. Quão disponíveis para você estão as informações que precisa no seu dia-a-dia?	0,506 (0,087)	-0,791 (0,164)
F211. Em que medida você tem oportunidades de atividade de lazer?	0,721 (0,100)	-0,120 (0,095)
F173. Quão satisfeito(a) você está com as condições do local onde mora?	0,639 (0,107)	-1,646 (0,237)
F193. Quão satisfeito(a) você está com o seu acesso aos serviços de saúde?	0,369 (0,074)	-0,832 (0,220)
F233. Quão satisfeito(a) você está com o seu meio de transporte?	0,517 (0,095)	-1,396 (0,240)
Questões gerais		
G1. Como você avaliaria sua qualidade de vida?	0,850 (0,120)	-0,959 (0,122)
G4. Quão satisfeito(a) você está com a sua saúde?	1,172 (0,156)	-0,032 (0,067)

TABELA 5 – Estatísticas clássicas para os itens do WHOQOL-BREVE

Item	Número de idosos satisfeitos (%)	Número de idosos insatisfeitos (%)	Correlação de Pearson	Correlação Bisserial
G1	293 (73,2)	107 (26,8)	0,472	0,635
G4	203 (50,8)	197 (49,2)	0,542	0,679
F14	230 (57,5)	170 (42,5)	0,461	0,582
F113	207 (51,8)	193 (48,2)	0,391	0,491
F21	252 (63,0)	148 (37,0)	0,542	0,693
F91	314 (78,5)	86 (21,5)	0,498	0,700
F33	241 (60,3)	159 (39,8)	0,279	0,354
F103	295 (73,8)	105 (26,2)	0,566	0,764
F124	255 (63,8)	145 (36,2)	0,554	0,710
F41	249 (62,2)	151 (37,8)	0,507	0,647
F242	337 (84,2)	63 (15,8)	0,414	0,626
F53	237 (59,2)	163 (40,8)	0,374	0,473
F71	319 (79,8)	81 (20,2)	0,307	0,438
F63	327 (81,8)	73 (18,2)	0,527	0,769
F81	224 (56,0)	176 (44,0)	0,461	0,580
F133	346 (86,5)	54 (13,5)	0,350	0,551
F153	148 (37,0)	252 (63,0)	0,241	0,308
F144	329 (82,2)	71 (17,8)	0,359	0,527
F161	265 (66,2)	135 (33,8)	0,385	0,499
F221	264(66,0)	136 (34,0)	0,399	0,516
F181	157 (39,2)	243 (60,8)	0,212	0,269
F201	257 (64,2)	143 (35,8)	0,362	0,465
F211	211 (52,7)	189 (47,2)	0,452	0,568
F173	325 (81,2)	75 (18,8)	0,367	0,532
F193	247 (61,8)	153 (38,2)	0,287	0,366
F233	297 (74,2)	103 (25,8)	0,329	0,446

Estatísticas descritivas dos escores brutos do WHOQOL-BREVE

	Global (26 itens)	Global (23 itens)
Média	17,075	15,695
Desvio-padrão	5,621	5,157
Mínimo	0	0
Máximo	26,0	23,0
Mediana	18,0	16,0

TABELA 6 - Matriz das correlações lineares de Pearson entre os escores latentes dos domínios e dos escores latentes globais (WHOQOL-BREVE)

Variável	Escore latente (26 itens)	Escore latente (23 itens)	Escore latente do domínio 1	Escore latente do domínio 2	Escore latente do domínio 3	Escore latente do domínio 4
Escore latente (26 itens)	1	0,975**	0,814**	0,748**	0,502**	0,405**
Escore latente (23 itens)		1	0,854**	0,787**	0,488**	0,422**
Escore latente do domínio 1			1	0,554**	0,351**	0,271**
Escore latente do domínio 2				1	0,399**	0,361**
Escore latente do domínio 3					1	0,712**
Escore latente do domínio 4						1

** Significativo a 1%.

Domínio 1 (físico)

Domínio 2 (psicológico)

Domínio 3 (relações sociais)

Domínio 4 (meio-ambiente)

TABELA 7 - Distribuição dos itens do Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS (WHOQOL-100), estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses

Elementos do WHOQOL-100 (itens)	Estimativas dos parâmetros	
	a (discriminação)	b (dificuldade)
Domínio I: físico		
F21. Você tem energia suficiente para seu dia-a-dia?	1,191 (0,132)	-0,450 (0,066)
F23. Quanto satisfeito(a) você está com a energia (disposição) que você tem?	1,533 (0,192)	-0,596 (0,061)
F31. Como você avaliaria o seu sono?	0,472 (0,079)	-0,596 (0,153)
F33. Quanto satisfeito(a) você está com o seu sono?	0,427 (0,076)	-0,628 (0,168)
Domínio II: psicológico		
F41. O quanto você aproveita a vida?	0,844 (0,106)	-0,492 (0,094)
F42. Em geral, você se sente contente?	1,395 (0,152)	-0,929 (0,088)
F43. Quanto otimista você se sente em relação ao futuro?	1,041 (0,122)	-0,528 (0,078)
F44. O quanto você experimenta sentimentos positivos em sua vida?	0,852 (0,105)	-0,900 (0,108)
F51. Como você avaliaria sua memória?	0,570 (0,082)	0,029 (0,114)
F52. Quanto satisfeito(a) você está com a sua capacidade de aprender novas informações?	0,698 (0,093)	-0,480 (0,101)
F53. O quanto você consegue se concentrar?	0,574 (0,083)	-0,454 (0,123)
F54. Quanto satisfeito(a) você está com sua capacidade de tomar decisões?	0,834 (0,111)	-1,005 (0,112)
F61. O quanto você se valoriza?	0,736 (0,110)	-1,431 (0,175)
F62. Quanta confiança você tem em si mesmo?	0,853 (0,127)	-1,705 (0,201)
F63. Quanto satisfeito(a) você está consigo mesmo?	1,406 (0,210)	-1,106 (0,092)
F64. Quanto satisfeito(a) você está com suas capacidades?	1,288 (0,176)	-0,858 (0,077)
F71. Você é capaz de aceitar sua aparência física?	0,581 (0,101)	-1,645 (0,240)
F74. Quanto satisfeito(a) você está com a aparência de seu corpo?	0,884 (0,121)	-0,938 (0,107)
Domínio III: nível de independência		
F91. Quanto bem você é capaz de se locomover?	1,169 (0,150)	-1,039 (0,095)
F92. Quanto satisfeito(a) você está com sua capacidade de se locomover?	1,005 (0,118)	-0,707 (0,091)
F101. Em que medida você é capaz de desempenhar suas atividades diárias?	1,300 (0,152)	-0,485 (0,064)

TABELA 7 - Distribuição dos itens do Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS (WHOQOL-100), estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses (continuação)

Elementos do WHOQOL-100 (itens)	Estimativas dos parâmetros	
	a (discriminação)	b (dificuldade)
F103. Quão satisfeito(a) você está com sua capacidade de desempenhar as atividades do seu dia-a-dia?	1,476 (0,178)	-0,764 (0,066)
F121. Você é capaz de trabalhar?	1,239 (0,133)	-0,329 (0,068)
F122. Você se sente capaz de fazer as suas tarefas?	1,423 (0,152)	-0,685 (0,065)
F123. Como você avaliaria a sua capacidade para o trabalho?	1,429 (0,159)	-0,386 (0,059)
F124. Quão satisfeito(a) você está com a sua capacidade para o trabalho?	1,696 (0,188)	-0,429 (0,052)
Domínio IV: relações sociais		
F132. Você se sente feliz com sua relação com as pessoas de sua família?	0,593 (0,128)	-2,573 (0,449)
F133. Quão satisfeito(a) você está com suas relações pessoais (amigos, parentes, conhecidos, colegas)?	0,778 (0,131)	-1,804 (0,222)
F134. Quão satisfeito(a) você está com sua capacidade de dar apoio aos outros?	1,029 (0,141)	-1,401 (0,132)
F141. Você consegue dos outros o apoio que necessita?	0,539 (0,087)	-1,179 (0,189)
F142. Em que medida você pode contar com amigos quando precisa deles?	0,565 (0,090)	-0,882 (0,148)
F143. Quão satisfeito(a) você está com o apoio que você recebe de sua família?	0,559 (0,110)	-2,078 (0,346)
F144. Quão satisfeito(a) você está com o apoio que você recebe de seus amigos?	0,719 (0,121)	-1,582 (0,203)
F151. Como você avaliaria sua vida sexual?	0,419 (0,075)	1,210 (0,244)
Domínio V: ambiente		
F161. Quão seguro(a) você se sente em sua vida diária?	0,620 (0,090)	-0,802 (0,139)
F162. Você acha que vive em um ambiente seguro?	0,419 (0,081)	-0,918 (0,209)
F171. Quão confortável é o lugar onde você mora?	0,456 (0,084)	-2,090 (0,359)
F172. Em que medida as características de seu lar correspondem às suas necessidades?	0,638 (0,107)	-1,364 (0,186)
F173. Quão satisfeito(a) você está com as condições do local onde mora?	0,674 (0,105)	-1,582 (0,210)
F174. O quanto você gosta de onde você mora?	0,705 (0,107)	-1,922 (0,0242)

TABELA 7 - Distribuição dos itens do Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS (WHOQOL-100), estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses (continuação)

Elementos do WHOQOL-100 (itens)	Estimativas dos parâmetros	
	a (discriminação)	b (dificuldade)
F183. Quão satisfeito(a) você está com sua situação financeira?	0,417 (0,077)	0,106 (0,152)
F192. Como você avaliaria a qualidade dos serviços de assistência social disponíveis para você?	0,518 (0,082)	-0,259 (0,125)
F193. Quão satisfeito(a) você está com o seu acesso aos serviços de saúde?	0,461 (0,066)	-0,697 (0,164)
F194. Quão satisfeito(a) você está com os serviços de assistência social?	0,650 (0,089)	-0,033 (0,102)
F201. Quão disponíveis para você estão as informações que precisa no seu dia-a-dia?	0,570 (0,090)	-0,728 (0,137)
F202. Em que medida você tem oportunidades de adquirir informações que considera necessárias?	0,593 (0,096)	-0,787 (0,136)
F203. Quão satisfeito(a) você está com as suas oportunidades de adquirir novas habilidades?	0,740 (0,092)	-0,169 (0,092)
F204. Quão satisfeito(a) você está com as suas oportunidades de obter novas informações?	0,740 (0,104)	-0,820 (0,113)
F211. Em que medida você tem oportunidades de atividade de lazer?	0,773 (0,101)	-0,125 (0,088)
F212. Quanto você é capaz de relaxar e curtir você mesmo?	0,955 (0,114)	-0,289 (0,077)
F213. O quanto você aproveita o seu tempo livre?	0,817 (0,108)	-0,864 (0,115)
F214. Quão satisfeito(a) você está com a maneira de usar o seu tempo livre?	0,874 (0,133)	-1,199 (0,140)
F221. Quão saudável é o seu ambiente físico (clima, barulho, poluição, atrativos)?	0,667 (0,093)	-0,740 (0,125)
F223. Quão satisfeito(a) você está com o seu ambiente físico (poluição, clima, barulho, atrativos)?	0,559 (0,088)	-0,778 (0,148)
F224. Quão satisfeito(a) você está com o clima do lugar em que vive?	0,491 (0,085)	-1,710 (0,281)
F231. Em que medida você tem meios de transporte adequados?	0,665 (0,092)	-0,693 (0,117)
F233. Quão satisfeito(a) você está com o seu meio de transporte?	0,589 (0,097)	-1,254 (0,192)

TABELA 7 - Distribuição dos itens do Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS (WHOQOL-100), estimativas dos parâmetros dos itens e respectivos erros-padrão entre parênteses (continuação)

Elementos do WHOQOL-100 (itens)	Estimativas dos parâmetros	
	a (discriminação)	b (dificuldade)
Domínio VI- Aspectos espirituais/Religião/Crenças pessoais		
F241. Suas crenças pessoais dão sentido à sua vida?	0,620 (0,110)	-1,946 (0,280)
F242. Em que medida você acha que sua vida tem sentido?	1,065 (0,155)	-1,378 (0,128)
F243. Em que medida suas crenças pessoais lhe dão força para enfrentar dificuldades?	0,679 (0,119)	-1,909 (0,262)
F244. Em que medida suas crenças pessoais lhe ajudam a entender as dificuldades da vida?	0,629 (0,111)	-1,677 (0,242)
Questões gerais		
G1. Como você avaliaria sua qualidade de vida?	0,973 (0,131)	-0,904 (0,105)
G2. Quão satisfeito(a) você está com a qualidade de sua vida?	1,407 (0,171)	-0,688 (0,071)
G3. Em geral, quão satisfeito(a) você está com a sua vida?	1,192 (0,143)	-0,835 (0,082)
G4. Quão satisfeito(a) você está com a sua saúde?	0,927 (0,107)	-0,037 (0,077)

TABELA 8 – Estatísticas clássicas para os itens do WHOQOL-100

Item	Número de idosos satisfeitos (%)	Número de idosos insatisfeitos (%)	Correlação de Pearson	Correlação Bisserial
F11	117 (29,2)	283 (70,8)	-0,184	-0,244
F12	154 (38,5)	246 (61,5)	-0,147	-0,187
F13	127 (31,8)	273 (68,2)	-0,161	-0,210
F14	89 (22,2)	311 (77,8)	-0,146	-0,204
F21	252 (63,0)	148 (37,0)	0,469	0,600
F22	122 (30,5)	278 (69,5)	-0,225	-0,296
F23	275 (68,8)	125 (31,2)	0,480	0,628
F24	107 (26,8)	293 (73,2)	-0,238	-0,320
F31	241 (60,3)	159 (39,7)	0,217	0,275
F32	107 (26,8)	293 (73,2)	-0,064	-0,086
F33	240 (60,0)	160 (40,0)	0,164	0,207
F34	95 (23,8)	305 (76,2)	-0,089	-0,123
F41	249 (62,2)	151 (37,8)	0,429	0,547
F42	310 (77,5)	90 (22,5)	0,492	0,684
F43	258 (64,5)	142 (35,5)	0,493	0,633
F44	288 (72,0)	112 (28,0)	0,464	0,619
F51	197 (49,2)	203 (50,8)	0,294	0,369
F52	243 (60,8)	157 (39,2)	0,381	0,485
F53	236 (59,0)	164 (41,0)	0,391	0,495
F54	296 (74,0)	104 (26,0)	0,442	0,597
F61	321 (80,2)	79 (19,8)	0,421	0,604
F62	346 (86,5)	54 (13,5)	0,377	0,594
F63	327 (81,8)	73 (18,2)	0,488	0,712
F64	300 (75,0)	100 (25,0)	0,461	0,628
F71	219 (79,8)	181 (20,2)	0,289	0,412
F72	44 (11,0)	356 (89,0)	-0,004	-0,007
F73	40 (10,0)	360 (90,0)	-0,021	-0,036
F74	293 (73,2)	107 (26,8)	0,428	0,575
F81	55 (13,8)	345 (86,2)	-0,283	-0,444
F82	147 (36,8)	253 (63,2)	-0,169	-0,216
F83	77 (19,2)	323 (80,8)	-0,279	-0,401
F84	90 (22,5)	310 (77,5)	-0,241	-0,336
F91	314 (78,5)	86 (21,5)	0,362	0,509
F92	276 (69,0)	124 (31,0)	0,352	0,461
F93	72 (18,0)	328 (82,0)	-0,234	-0,342
F94	69 (17,2)	331 (82,8)	-0,153	-0,226
F101	258 (64,5)	142 (35,5)	0,469	0,603
F102	67 (16,8)	333 (83,2)	-0,241	-0,359
F103	294 (73,5)	106 (26,5)	0,457	0,616
F104	79 (19,8)	321 (80,2)	-0,266	-0,381
F111	130 (32,5)	270 (67,5)	-0,048	-0,063
F112	124 (31,0)	276 (69,0)	-0,044	-0,058
F113	120 (30,0)	280 (70,0)	-0,063	-0,084

TABELA 8 – Estatísticas clássicas para os itens do WHOQOL-100 (continuação)

Item	Número de idosos satisfeitos (%)	Número de idosos insatisfeitos (%)	Correlação de Pearson	Correlação Bisserial
F114	127 (31,8)	273 (68,2)	-0,109	-0,142
F121	238 (59,5)	162 (40,5)	0,498	0,631
F122	284 (71,0)	116 (29,0)	0,494	0,655
F123	247 (61,8)	153 (39,2)	0,485	0,617
F124	255 (63,8)	145 (36,2)	0,535	0,685
F131	63 (15,8)	337 (84,2)	-0,191	-0,289
F132	361 (90,2)	39 (9,8)	0,288	0,496
F133	346 (86,5)	54 (13,5)	0,397	0,625
F134	337 (84,2)	63 (15,8)	0,413	0,625
F141	286 (71,5)	114 (28,5)	0,385	0,512
F142	268 (67,0)	132 (33,0)	0,372	0,483
F143	338 (84,5)	62 (15,5)	0,308	0,468
F144	329 (82,2)	71 (17,8)	0,379	0,556
F151	126 (31,5)	274 (68,5)	0,270	0,353
F152	110 (27,5)	290 (72,5)	0,268	0,358
F153	147 (36,8)	253 (63,2)	0,235	0,300
F154	36 (9,0)	364 (91,0)	-0,036	-0,063
F161	266 (66,5)	134 (33,5)	0,387	0,502
F162	257 (64,2)	143 (35,8)	0,289	0,371
F163	230 (57,5)	170 (42,5)	-0,045	-0,056
F164	193 (48,2)	207 (51,8)	0,257	0,322
F171	324 (81,0)	76 (19,0)	0,304	0,440
F172	308 (77,0)	92 (23,0)	0,370	0,513
F173	325 (81,2)	75 (18,8)	0,361	0,523
F174	346 (86,5)	54 (13,5)	0,321	0,505
F181	157 (39,2)	243 (60,8)	0,220	0,279
F182	87 (21,8)	313 (78,2)	-0,201	-0,282
F183	193 (48,2)	207 (51,8)	0,260	0,326
F184	142 (35,5)	258 (64,5)	-0,199	-0,255
F191	247 (61,8)	153 (38,2)	0,319	0,406
F192	219 (54,8)	181 (45,2)	0,347	0,436
F193	247 (61,8)	153 (38,2)	0,354	0,452
F194	202 (50,5)	198 (49,5)	0,438	0,548
F201	257 (64,2)	143 (35,8)	0,373	0,479
F202	263 (65,8)	237 (34,2)	0,382	0,493
F203	215 (53,8)	185 (46,2)	0,431	0,541
F204	275 (68,8)	125 (31,2)	0,457	0,598
F211	211 (52,7)	189 (47,3)	0,452	0,567
F212	230 (57,5)	170 (42,5)	0,541	0,682
F213	283 (70,8)	117 (29,2)	0,455	0,602
F214	314 (78,5)	86 (21,5)	0,437	0,614
F221	264 (66,0)	136 (34,0)	0,408	0,527
F222	74 (18,5)	326 (81,5)	-0,024	-0,035
F223	260 (65,0)	140 (35,0)	0,364	0,469

TABELA 8 – Estatísticas clássicas para os itens do WHOQOL-100 (continuação)

Item	Número de idosos satisfeitos (%)	Número de idosos insatisfeitos (%)	Correlação de Pearson	Correlação Bisserial
F224	311 (77,8)	89 (22,2)	0,296	0,413
F231	260 (65,0)	140 (35,0)	0,406	0,523
F232	57 (14,2)	343 (85,8)	-0,148	-0,229
F233	296 (74,0)	104 (26,0)	0,298	0,403
F234	52 (13,0)	348 (87,0)	-0,209	-0,332
F241	339 (84,8)	61 (15,2)	0,309	0,471
F242	337 (84,2)	63 (15,8)	0,403	0,609
F243	343 (85,8)	57 (14,2)	0,340	0,527
F244	326 (81,5)	74 (18,5)	0,324	0,471
G1	294 (73,5)	106 (26,5)	0,392	0,529
G2	284 (71,0)	116 (29,0)	0,508	0,674
G3	295 (73,8)	105 (26,2)	0,458	0,618
G4	202 (50,5)	198 (49,5)	0,357	0,448

Estatísticas descritivas do escore bruto do WHOQOL-100		
	Global (100 itens)	Global (65 itens)
Média	54,622	44,630
Desvio-padrão	11,566	13,285
Mínimo	11,0	1,0
Máximo	79,0	64,0
Mediana	57,0	47,0

TABELA 9 - Matriz das correlações lineares de Pearson entre os escores latentes dos domínios e dos escores latentes globais (WHOQOL-100)

Variável	E.L. global 65 itens	E.L. domínio 1	E.L. domínio 2	E.L. domínio 3	E.L. domínio 4	E.L. domínio 5	E.L. domínio 6	E.L. fac geral
E.L. global 65 itens	1	-0,511**	0,845**	0,725**	0,497**	-0,297**	-0,159**	-0,370**
E.L. domínio 1		1	-0,464**	-0,499**	-0,145**	0,116**	0,068	0,193**
E.L. domínio 2			1	0,587**	0,445**	-0,296**	-0,142**	-0,334**
E.L. domínio 3				1	0,290**	-0,155**	-0,101*	-0,364**
E.L. domínio 4					1	-0,646**	-0,130**	-0,120*
E.L. domínio 5						1	0,051	0,023
E.L. domínio 6							1	0,118**
E.L. fac geral								1

** Significativo a 1%.

* Significativo a 5%.

E.L: escore latente

Domínio 1 (físico)

Domínio 2 (psicológico)

Domínio 3 (nível de independência)

Domínio 4 (social)

Domínio 5 (meio-ambiente)

Domínio 6 (espiritual)

7. CONCLUSÕES E CONSIDERAÇÕES FINAIS

Em relação à Escala de Qualidade de Vida de Flanagan (EQVF), pode-se dizer que a escala apresentou alta fidedignidade, tomada globalmente; e pela TRI, o item 5 (vida íntima) mostrou-se inadequado por apresentar uma baixa discriminação, devendo ser retirado do instrumento. Os idosos ativos não estão satisfeitos com a sua aprendizagem, apresentando um valor alto para o parâmetro de dificuldade. Os itens relativos ao bom nível de autoconhecimento, emprego, boa comunicação criativa e oportunidades de lazer são bons indicadores de níveis superiores de qualidade de vida. As variáveis níveis de depressão e saúde percebida são significantes para o escore latente, obtido via TRI, acrescidas da variável nível de desesperança, significativa apenas para o escore bruto.

As estatísticas clássicas comprovaram os resultados da TRI, indicando que a maioria dos idosos estão contentes com a sua família, bem como com assistir TV ou cinema. Os escores latente e bruto mantêm uma alta correlação entre si, e a classificação dos idosos a partir da posição da mediana concordam nas duas representações de escores.

A estrutura apresentada na AF não é a mesma da escala original, com 5 ao invés de 3 fatores. As cargas fatoriais apresentaram uma relação com as estimativas dos parâmetros de inclinação, e os itens mais discriminantes, na TRI, coincidem com os itens com maiores cargas fatoriais do primeiro fator na AF.

Com o WHOQOL-BREVE, que apresentou uma boa fidedignidade global e por domínios, pôde-se observar que a estrutura encontrada nas análises não foi a mesma da proposta pelo Grupo WHOQOL, com 5 domínios, ao invés dos 4. Com base na TRI, 3 dos 26 itens devem ser retirados devido à inconsistência. Os itens são referentes à vida sexual (item constrangedor), dinheiro suficiente para satisfazer suas necessidades e satisfação com o acesso aos serviços de saúde. Os idosos mostraram-se satisfeitos com sua qualidade de vida em geral.

Os fatores influentes nos escores latentes e brutos são os mesmos: saúde percebida, níveis de depressão e desesperança. Os escores, bruto e latente, apresentaram alta correlação entre si, bem como os escores latentes por domínios com o escore latente global. Os idosos classificados pelos seus escores, tanto brutos,

quanto latentes, tiveram concordância na classificação pela mediana, indicando que usar o escore latente, ou bruto, não influencia na categorização de alta ou baixa qualidade de vida.

Os resultados da análise clássica confirmaram os achados pela TRI, onde percebeu-se que os idosos estão satisfeitos com suas relações sociais e insatisfeitos com sua vida íntima.

No WHOQOL-100, foram encontrados 35 itens inconsistentes, o que fez com que o instrumento perdesse a sua caracterização original. O instrumento, mesmo com tantos itens retirados, seja globalmente ou por domínios, mostrou-se fidedigno.

A estrutura da AF apresentou um fator a menos que a proposta pela OMS, e também os itens mais discriminantes estavam no primeiro fator e apresentando valores para as cargas fatoriais semelhantes às estimativas dos parâmetros de inclinação dos itens. Mesmo apresentando muitos itens inconsistentes, que foram retirados dos cálculos das habilidades finais, os restantes apresentaram um bom poder de discriminação. Os idosos relataram baixos níveis de satisfação com a sua memória, vida sexual e situação financeira.

Os idosos foram classificados de forma semelhante, seja ao serem considerados os escores latentes ou brutos, com base na mediana. Os fatores influentes nos escores brutos e latentes, que apresentaram intensa correlação, são os mesmos apresentados no WHOQOL-BREVE. O escore latente global apresentou correlações negativas com os domínios físico, ambiente e espiritual, e correlações positivas com os domínios psicológico, nível de independência e social. Os escores latentes do WHOQOL-BREVE e WHOQOL-100 apresentaram uma alta correlação, indicando que ao serem considerados de forma global, o WHOQOL-BREVE é uma boa alternativa para o instrumento maior. Nas correlações entre os escores dos domínios do WHOQOL-BREVE e WHOQOL-100, foram encontrados resultados conflitantes. Para o domínio físico e ambiente, os escores apresentaram uma correlação negativa entre os instrumentos, e para os domínios psicológico e social, as correlações foram positivas. Esses dois domínios (psicológico e social) estão medindo a qualidade de vida de forma equivalente e na mesma direção, ao passo que os outros dois não.

As interpretações referentes ao instrumento WHOQOL-100 devem merecer

uma atenção especial. A retirada de muitos itens, necessária após verificação de inconsistência na TRI, produziu resultados divergentes. Outros aspectos importantes são a dicotomização das informações, que por um lado, resulta na perda de informações, mas por outro tornou possível a aplicação do modelo, já que o WHOQOL-100 não ajustou nenhum modelo para resposta politômica, e tampouco seria possível a estimação de tantos parâmetros com o tamanho amostral disponível. Os modelos para respostas dicotômicas têm outro aspecto favorável: a interpretação dos parâmetros, que são relativamente mais simples e de interpretação mais clara que os demais, considerados mais complexos. Uma outra característica que merece atenção especial é a dimensionalidade das escalas. Mesmo sabendo da multidimensionalidade, são empregados modelos unidimensionais na avaliação de qualidade de vida. Há de se atentar também para a grande porcentagem de mulheres presente no estudo, ainda que ao verificar comportamento diferencial entre os itens para os grupos, o mesmo não tenha ocorrido. E ao fato do WHOQOL-BREVE ter sido construído com base nas respostas ao instrumento com 100 itens, podendo ocorrer diferença nas respostas caso fosse aplicado diretamente. Outro aspecto importante são os itens com 2 pólos nos instrumentos da OMS. Esses itens medem a qualidade de vida, mas em direções contrárias, podendo comprometer em parte os resultados encontrados.

De forma geral é possível concluir que a TRI constituiu-se numa ferramenta promissora e útil na análise de observações geradas por instrumentos que visam avaliar qualidade de vida, especialmente em idosos. Mesmo que os resultados das análises pelos métodos clássicos tenham sido muito semelhantes aos do modelo de TRI, a mesma é preferível aos anteriores, visto que aproveita melhor a informação presente em cada item, diferenciando os itens mais importantes dos menos importantes, além da flexibilidade de lidar com dados faltantes, o que não compromete a aplicação do modelo de resposta ao item, visto que o aspecto central são os itens e não o instrumento como um todo.

Em suma, recomenda-se usar os modelos de TRI, ainda que esses modelos sejam muito mais complexos que os demais. A TRI apresenta uma série de vantagens, constituindo-se em mecanismo promissor na solução de problemas vinculados à avaliação de qualidade de vida. Torna-se necessária uma maior difusão

e disseminação desta técnica. Este trabalho não visou abordar todos os assuntos vinculados à TRI, nem tampouco pretendeu esgotar o tema, mas alertar para a importância e eficiência da TRI, especialmente na avaliação de qualidade de vida.

8. ANEXOS

Anexo A - Projeto de Pesquisa

Anexo B - Aprovação pela Comissão de Pesquisa

Anexo C – Autorização do responsável para utilizar a base de dados

Anexo D – Escala de Qualidade de Vida de Flanagan

Anexo E – Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS (WHOQOL-100)

Anexo F – Instrumento de Avaliação de Qualidade de Vida desenvolvido pela OMS – Versão abreviada (WHOQOL-BREVE)

Anexo G – Estimativas das habilidades (traços latentes) dos idosos

Anexo H – Tabelas que não foram colocadas no artigo

Anexo I – Resultados adicionais que não foram colocados no artigo

Anexo J – Estatísticas de concordância

Anexo K – Recursos Computacionais

Anexo A - Projeto de Pesquisa

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA

PROJETO DE PESQUISA

TEORIA DE RESPOSTA AO ITEM APLICADA À EPIDEMIOLOGIA

AUTOR: KATRINE GUEWEHR
ORIENTADOR: JOÃO RIBOLDI

Porto Alegre, abril de 2006.

1. INTRODUÇÃO

1.1 Questão de Pesquisa

A avaliação de fenômenos epidemiológicos que envolvem qualidade de vida relacionada à saúde tem sido comum, porém tratada analiticamente de forma não suficientemente satisfatória e consistente. A Teoria de Resposta ao Item (TRI) trata desfechos desta natureza com consistência e eficiência e tem se constituído em mecanismo promissor para a solução do problema.

A TRI, muito embora tenha sido introduzida na década de 50, somente nos últimos 15 anos é que vem sendo aplicada em larga escala, principalmente em avaliações educacionais (1). Recentemente passou a ser aplicada em Medicina, especialmente na avaliação de qualidade de vida.

A TRI é uma reunião de modelos estatísticos usados para fazer predições e estimativas sobre as habilidades, competências e mensurações de qualidade de vida, medidas a partir de um instrumento. Essa técnica propõe modelos para os traços latentes, ou seja, características que não podem ser observadas diretamente, sendo inferidas através da observação de variáveis secundárias que estejam relacionadas a ela (1).

A metodologia de TRI sugere formas de representar a relação entre a probabilidade de um indivíduo dar uma certa resposta a um item, por exemplo, um questionário de qualidade de vida, e seus traços latentes da área avaliada. O modelo permite obter, para cada sujeito, um único valor para a estrutura latente, o qual resume a informação contida nas suas respostas aos vários itens com o mínimo de perda de informação. A metodologia é sofisticada e complexa, tornando-se possível apenas com recursos computacionais especializados. Apresenta também um amplo espectro de vinculação com outros procedimentos estatísticos de análise de dados, como modelos de análise fatorial, análise de sobrevivência e análise de medidas repetidas.

Este trabalho tem por objetivo principal modelar dados de qualidade de vida usando a Teoria de Resposta ao Item, visto que a TRI utiliza potencialmente de

forma mais eficiente a informação disponível que as outras técnicas mais simples, viabilizando a redução do número de pacientes nos estudos sem redução de poder.

1.2 Teoria de Resposta ao Item

A TRI surgiu com o intuito de solucionar uma série de limitações que a Teoria Clássica dos Testes (TCT) ou Teoria Clássica das Medidas (TCM) apresentava. Os modelos de TRI têm vantagens distintas sobre os métodos clássicos: o fato dos parâmetros dos itens e níveis do traço latente serem independentes; o modelo ser expresso no nível da resposta do item observado além do nível do escore do teste observado; a contribuição de cada item para a escala pode ser medida através da informação de TRI e a criação de formas curtas, paralelas e customizadas para as escalas, além de outras não tão importantes (2).

Os modelos de TRI são funções de um ou mais parâmetros dos itens, dependendo da complexidade do modelo. De uma forma geral esses parâmetros caracterizam dificuldade e discriminação dos itens, e são estimados, preferentemente, pelos métodos de máxima verossimilhança ou por métodos bayesianos, como o MCMC (Markov Chains Monte Carlo Methods), que utilizam procedimentos iterativos não triviais e exigem intenso envolvimento computacional. Os modelos de TRI dependem fundamentalmente de três fatores: da natureza do item (dicotômicos ou politômicos), do número de populações envolvidas (uma ou mais) e da quantidade de traços latentes (um ou mais). Os modelos tradicionalmente utilizados para a análise de dados de qualidade de vida utilizam itens com respostas politômicas. Existem quatro tipos de modelos de TRI para respostas politômicas que são candidatos naturais a modelar dados de qualidade de vida: Modelo de Escala Gradual de Andrich, Modelo de Resposta Gradual de Samejima, Modelo de Crédito Parcial de Masters e Modelo de Crédito Parcial Generalizado de Muraki.

Maiores detalhes sobre TRI encontram-se em (1;3).

1.3 TRI e Qualidade de Vida

A Organização Mundial de Saúde define saúde como “um estado de bem-estar completo físico, mental e social e não meramente a ausência de doença”. A Qualidade de vida relacionada à saúde (HRQOL) irá medir esses 3 ou 4 componentes.

Há inúmeros instrumentos que abordam a questão de qualidade de vida relacionada à saúde. Dentre eles pode-se destacar o ADL, que trata das atividades diárias relacionadas à saúde em geral, ESSI, um instrumento com 7 itens relacionados ao suporte social, EORTC QLQ C30, um instrumento com 30 itens para estudos de câncer, FACT-An, uma sub-escala de 13 itens do FACT para anemia, HAQ, um questionário de medida de saúde usado em estudos de artrite, PAR-Q, um questionário relacionado às atividades físicas, SDS, uma escala de 18 itens para sintomas de aflição e o SF-36, um questionário de 36 itens relacionados à saúde em geral (4).

Muitas dimensões de qualidade de vida podem ser medidas, como saúde geral, funcionamento físico, social, saúde mental, dentre outros, e os instrumentos referentes ao aspecto de qualidade de vida selecionado devem ser validados dentro do contexto em que ele será usado. Por exemplo, uma escala que tem se mostrada válida para pacientes com doenças cardiovasculares pode não ser apropriada para pacientes com câncer.

Um dos primeiros estudos relacionando TRI e qualidade de vida relacionada à saúde foi um estudo comparativo entre os modelos de TRI, no caso, o modelo de Rasch e o método tradicional de soma de escores, aplicadas em uma escala de funcionamento físico com 10 itens (PF-10), utilizando-se a precisão relativa (5).

Outra grande e recente contribuição da TRI é a criação dos testes adaptativos computadorizados (CAT), que escolhem questões de um banco de itens através de mecanismos probabilísticos, e cujo objetivo é maximizar a precisão com um número mínimo de questões. Um exemplo de CAT é a pesquisa da avaliação do impacto combinado do HIV e seu tratamento sobre o funcionamento e bem estar na perspectiva do paciente, utilizando TRI, banco de itens e CAT (6).

A grande maioria dos trabalhos publicados relacionando TRI e qualidade de vida relacionada à saúde utilizam o modelo de resposta gradual e o modelo de crédito parcial generalizado. Em Uttaro et. al (1999) foi utilizado o modelo de resposta gradual com o objetivo de extrair informação útil visando melhorar a qualidade de vida de pacientes com doença mental severa. O instrumento utilizava uma escala de Likert com 7 alternativas divididas em vários construtos (2).

Inúmeros trabalhos na literatura utilizaram o modelo de crédito parcial generalizado. Dentre estes destaca-se Bjorner et al., 2003, que avaliaram o impacto da dor de cabeça e enxaqueca através do instrumento HIT (Headache Impact Test) e testes adaptativos computadorizados, chegando-se ao final com uma versão customizada com 6 itens e auto-administrada (7).

Em outro estudo em 2003, com um instrumento construído para avaliar fadiga (FACIT-F), utilizou-se na análise uma extensão politômica do modelo dicotômico de Rasch, na avaliação de resultados de 2 grupos: um grupo de pacientes com câncer e o outro da população em geral (8).

Para um dos questionários mais usados nesses últimos anos, EORTC QLQ-C30 (European Organization for Research and Treatment of Cancer Quality of Life Group), foi publicada uma versão abreviada por Bjorner et. al. em 2004, próprio para pesquisas em cuidado paliativo. Analisou-se adicionalmente 2 escalas menores, apresentando ótima concordância com o escore total, além de pouco viés. A modelagem foi procedida utilizando-se o modelo de crédito parcial generalizado (9). Com base nesse mesmo instrumento, Tejada & Rojas em 2005, utilizaram um modelo de escala gradual para estimar uma medida de qualidade de vida de pacientes terminais de câncer (10). E em outro estudo, realizado por Petersen et al., em 2005, foi apresentado um estudo comparativo entre os modelos de TRI e o método tradicional de soma de escores, aplicados ao EORTC QLQ-C30, através da validade relativa, sendo analisadas três escalas maiores: funcionamento físico, emocional e fadiga e outras 6 escalas menores. Utilizou-se o modelo de crédito parcial generalizado (11).

2. OBJETIVOS

2.1 Objetivo Geral:

O objetivo geral deste trabalho é estruturar metodologicamente a Teoria de Resposta ao Item, visando sua aplicação na solução de problemas epidemiológicos que envolvam qualidade de vida e estabelecer sua vinculação com outras técnicas estatísticas.

2.2 Objetivos Específicos:

- 1) Aplicar TRI em dados de qualidade de vida e comparar com os resultados gerados pela Teoria Clássica das Medidas.

- 2) Comparar os resultados gerados pela TRI com aqueles gerados pela Análise Fatorial.

3. METODOLOGIA

3.1 Caracterização do estudo

Os dados utilizados para a análise são referentes a um estudo observacional e metodológico, realizado por Moraes (2004), no período de junho de 2000 a maio de 2002. O período de coleta dos dados foi entre maio e dezembro de 2002. O delineamento foi do tipo transversal. Na amostra, foram incluídos, por conveniência, 400 sujeitos socialmente ativos, com 60 anos ou mais (12).

3.1.1 Coleta dos dados

Os dados foram coletados a partir de informações dadas pelos voluntários através de entrevistas estruturadas, que foram realizadas com a ajuda de acadêmicos treinados para tal tarefa.

Os idosos após a assinatura do termo de consentimento livre e esclarecido, responderam a um conjunto de itens, onde primeiramente foram perguntados dados como sexo, idade, data de nascimento, nível educacional, estado civil, como está sua saúde e problemas de saúde atual ou condição presente. Após essa etapa inicial foram utilizados como instrumentos de pesquisa a Escala de Qualidade de Vida de Flanagan e a Avaliação de Qualidade de Vida elaborada pela OMS-WHOQOL-100, sendo avaliados aspectos relacionados às propriedades psicométricas. Também foram utilizadas a Escala de Depressão Geriátrica com 15 itens e a Escala de Desesperança de Beck com 20 itens.

Neste trabalho serão utilizados os dados coletados através da Escala de Qualidade de Vida de Flanagan (EQVF) e do instrumento de pesquisa WHOQOL-100. Essa primeira escala, que foi traduzida e validada no Brasil, possui 15 itens divididos em 5 dimensões ou construtos: bem estar físico e mental, relações com outras pessoas, atividades sociais, comunitárias e cívicas, desenvolvimento pessoal e realização e recreação, avaliadas por uma escala com 7 categorias (12).

O WHOQOL-100 também foi traduzido e validado no Brasil, sendo composto por 6 domínios: físico, psicológico, nível de independência, relações

sociais, ambiente e espiritualidade/crenças/religiões, distribuídos em uma escala com 5 categorias. A Organização Mundial de Saúde desenvolveu este instrumento com uma perspectiva transcultural, e o mesmo já se encontra disponível em 20 idiomas diferentes (12).

3.2 Análise Estatística

Para a modelagem dos dados de qualidade de vida serão utilizados os quatro modelos candidatos para respostas politômicas: Modelo de Escala Gradual de Andrich, Modelo de Crédito Parcial de Masters, Modelo de Resposta Gradual de Samejima e Modelo de Crédito Parcial Generalizado de Muraki. Os dois primeiros modelos utilizam apenas o parâmetro de dificuldade (modelos de 1 parâmetro), e os outros 2, incorporam o parâmetro de discriminação (modelos de 2 parâmetros). Comparar-se-á os modelos e adotar-se-á o modelo de melhor performance. Com o modelo selecionado comparar-se-á seus resultados com aqueles gerados pela Teoria Clássica das Medidas e pela Análise Fatorial.

As análises utilizando-se os modelos de TRI serão procedidas no software PARSCALE. Para a TCM e a análise fatorial utilizar-se-á o SAS e o SPSS.

4. ASPECTOS ÉTICOS

Este trabalho utiliza como fonte de dados aqueles coletados por Moraes, 2004 (12). Todos os pacientes incluídos na pesquisa assinaram o Termo de Consentimento Informado, além de outros detalhes éticos formais necessários.

6. REFERÊNCIAS

- (1) Andrade DF, Tavares HR, Valle RC. Teoria de Resposta do Item: Conceitos e Aplicações. SINAPE (Simpósio Nacional de Probabilidade e Estatística). 2000.
- (2) Uttaro T, Lehman A. Graded response modeling of the Quality of Life Interview. Evaluation and Program Planning. 22 ed. 1999. p. 41-52.
- (3) Valle RC. Teoria de Resposta ao Item. São Paulo: USP, 99 f. Dissertação (Mestrado em estatística) – Universidade de São Paulo 1999.
- (4) Gnecco C, Lachenbruch PA. Regulatory Aspects of Quality of Life. Statistical Methods for Quality of Life Studies. Kluwer Academic Publishers; 2002. p. 9-19.
- (5) McHorney CA, Haley SM, Ware JE, Jr. Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. J Clin Epidemiol 1997 Apr;50(4):451-61.
- (6) Revicki DA, Cella DF. Health status assessment for the twenty-first century: item response theory, item banking and computer adaptive testing. Qual Life Res 1997 Aug;6(6):595-600.
- (7) Bjorner JB, Kosinski M, Ware JE, Jr. Using item response theory to calibrate the Headache Impact Test (HIT) to the metric of traditional headache scales. Qual Life Res 2003 Dec;12(8):981-1002.
- (8) Lai JS, Cella D, Chang CH, Bode RK, Heinemann AW. Item banking to improve, shorten and computerize self-reported fatigue: an illustration of steps to create a core item bank from the FACIT-Fatigue Scale. Qual Life Res 2003 Aug;12(5):485-501.
- (9) Bjorner JB, Petersen MA, Groenvold M, Aaronson N, Ahlner-Elmqvist M, Arraras JI, et al. Use of item response theory to develop a shortened version of the EORTC QLQ-C30 emotional functioning scale. Qual Life Res 2004 Dec;13(10):1683-97.
- (10) Rojas AJ, Rojas OM. Application of an IRT polytomous model for measuring Health Related Quality of Life. Social Indicators Research 2005.
- (11) Petersen MA, Groenvold M, Aaronson N, Brenne E, Fayers P, Nielsen JD, et al. Scoring based on item response theory did not alter the measurement ability of EORTC QLQ-C30 scales. J Clin Epidemiol 2005 Sep;58(9):902-8.
- (12) Moraes JF. Fatores determinantes do envelhecimento bem-sucedido do idoso socialmente ativo da região metropolitana de Porto Alegre. Porto Alegre: PUC –RS, Tese (Dourorado em Gerontologia Médica) – Pontifícia Universidade Católica do Rio Grande do Sul 2004.

Anexo B - Aprovação pela Comissão de Pesquisa

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA

COMISSÃO DE PESQUISA

PARECER

RELATO

Trata-se de projeto de pesquisa intitulado "Teoria De Resposta Ao Item Aplicada À Epidemiologia" que será coordenado pela Prof. Dr. João Riboldi, na orientação da aluna, do PPG em Epidemiologia desta Universidade, KatrineGuewerh.

ANÁLISE

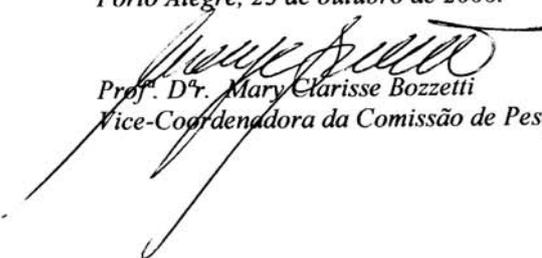
O projeto tem o objetivo estruturar metodologicamente a Teoria de Resposta ao Item, visando sua aplicação na solução de problemas epidemiológicos que envolvam qualidade de vida e estabelecer sua vinculação com outras técnicas estatísticas. Os dados utilizados para a análise são referentes a um estudo observacional realizado por Moraes (2004), no período de junho de 2000 a maio de 2002. Este está adequado metodologicamente, incluindo a proposta de análise estatística.

A proposta envolve um método de análise com dados de uma base de dados uma tese de doutorado já defendida de Moraes JF. **Fatores determinantes do envelhecimento bem-sucedido do idoso socialmente ativo da região metropolitana de Porto Alegre. Porto Alegre: PUC –RS, Tese (Doutorado em Gerontologia Médica) – Pontifícia Universidade Católica do Rio Grande do Sul 2004.** Como o proponente não coletará dados e, sim, utilizará um banco de dados já existente, cabe ao mesmo apresentar uma carta dos responsáveis por este banco indicando que os mesmos estão de acordo com a utilização das informações desta base para esta análise específica. Este documento deverá ser anexado ao projeto.

CONCLUSÃO

Com a inclusão deste documento solicitado por esta Comissão o projeto estará adequado para execução, sem necessidade de encaminhamento à Comissão de Ética desta Universidade.

Porto Alegre, 23 de outubro de 2006.


Prof.ª D^{ra}. Mary Clarisse Bozzetti
Vice-Coordenadora da Comissão de Pesquisa

Anexo C – Autorização do responsável para utilizar a base de dados

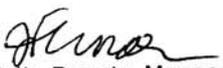
Porto Alegre, 08 de dezembro de 2006.

Ao Prof. Dr. João Riboldi
Orientador – PPG Epidemiologia da UFGRS

Eu, Dr. João Feliz Duarte de Moraes, professor adjunto dos Departamentos de Estatística da UFGRS e da PUCRS, autorizo a acadêmica Katrine Guewehr, aluna do PPG em Epidemiologia, a utilizar as informações do Banco de Dados referente à Qualidade de Vida de idosos socialmente ativos da região Metropolitana de Porto Alegre, coletados durante a elaboração de minha tese de Doutorado em Gerontologia Biomédica, realizada na PUCRS e com a aprovação do CEP-PUCRS, sob nº 01/01065, para o desenvolvimento de seu projeto "Teoria da Resposta ao Item Aplicada à Epidemiologia".

Atenciosamente,

Prof. Dr. João Feliz Duarte de Moraes


João Feliz Duarte Moraes
Estatístico 99491776
jfmoraes@puccrs.br

Anexo D – Escala de Qualidade de Vida de Flanagan

Escala de Qualidade de Vida de Flanagan – EQVF

A escala EQVF busca avaliar a qualidade de vida utilizando as seguintes expressões linguísticas :

Muito Insatisfeito	Insatisfeito	Pouco Satisfeito	Indiferente	Pouco Satisfeito	Satisfeito	Muito Satisfeito
1	2	3	4	5	6	7

As expressões linguísticas são atribuídos escores numa faixa de 1 a 7 pontos, conforme indicado acima. Responda cada um dos itens abaixo assinalando o escore que indica seu grau de satisfação em relação aos seguintes aspectos de sua vida:

Qual a sua satisfação em relação a:	1	2	3	4	5	6	7
1. Conforto material: casa, alimentação, situação financeira.							
2. Saúde: fisicamente bem e vigoroso(a).							
3. Relacionamento com pais, irmãos e outros parentes: comunicação, visita e ajuda.							
4. Construir família: ter e criar filhos.							
5. Relacionamento íntimo com esposo(a), namorado(a) ou outra pessoa relevante.							
6. Amigos próximos: compartilhar interesses, atividades e opiniões.							
7. Voluntariamente, ajudar e apoiar outras pessoas.							
8. Participação em associações e atividades de interesse público.							
9. Aprendizagem: freqüentar outros cursos para conhecimentos gerais.							
10. Auto-conhecimento: reconhecer seus potenciais e limitações.							
11. Trabalho (emprego ou em casa): atividade interessante, gratificante que vale a pena.							
12. Comunicação criativa.							
13. Participação em recreação ativa.							
14. Ouvir música, assistir TV ou cinema, leitura ou outros entretenimentos.							
15. Socialização: “fazer amigos”.							

Utilize o verso desta folha, se desejar, para falar sobre QUALIDADE DE VIDA PARA VOCÊ:.....

Dimensões da escala de Flanagan	
Dimensões da EQVF	Itens
1. Bem estar físico e material.	1 e 2
2. Relações com outras pessoas.	3,4,5 e 6
3. Atividades sociais, comunitárias e cívicas.	7 e 8
4. Desenvolvimento pessoal e realização.	9,10,11 e 12
5. Recreação.	13, 14 e 15

**Anexo E – Instrumento de Avaliação de Qualidade de Vida desenvolvido pela
OMS (WHOQOL-100)**

WHOQOL-100

Por favor, leia cada questão, veja o que você acha e circule no número e lhe parece a melhor resposta.

nada	muito pouco	mais ou menos	bastante	extremamente
1	2	3	4	5

F1.2 Você se preocupa com sua dor ou desconforto (físicos)?	1	2	3	4	5
F1.3 Quão difícil é para você lidar com alguma dor ou desconforto?	1	2	3	4	5
F1.4 Em que medida você acha que sua dor (física) impede você de fazer o que você precisa?	1	2	3	4	5
F2.2 Quão facilmente você fica cansado(a)?	1	2	3	4	5
F2.4 O quanto você se sente incomodado(a) pelo cansaço?	1	2	3	4	5
F3.2 Você tem alguma dificuldade para dormir (com o sono)?	1	2	3	4	5
F3.4 O quanto algum problema com o sono lhe preocupa?	1	2	3	4	5
F4.1 O quanto você aproveita a vida?	1	2	3	4	5
F4.3 Quão otimista você se sente em relação ao futuro?	1	2	3	4	5
F4.4 O quanto você experimenta sentimentos positivos em sua vida?	1	2	3	4	5
F5.3 O quanto você consegue se concentrar?	1	2	3	4	5
F6.1 O quanto você se valoriza?	1	2	3	4	5
F6.2 Quanta confiança você tem em si mesmo?	1	2	3	4	5
F7.2 Você se sente inibido(a) por sua aparência?	1	2	3	4	5
F7.3 Há alguma coisa em sua aparência que faz você não se sentir bem?	1	2	3	4	5
F8.2 Quão preocupado(a) você se sente?	1	2	3	4	5
F8.3 Quanto algum sentimento de tristeza ou depressão interfere no seu dia-a-dia?	1	2	3	4	5
F8.4 O quanto algum sentimento de depressão lhe incomoda?	1	2	3	4	5
F10.2 Em que medida você tem dificuldade em exercer suas atividades do dia-a-dia?	1	2	3	4	5
F10.4 Quanto você se sente incomodado por alguma dificuldade em exercer as atividades do dia-a-dia?	1	2	3	4	5
F11.2 Quanto você precisa de medicação para levar a sua vida do dia-a-dia?	1	2	3	4	5
F11.3 Quanto você precisa de algum tratamento médico para levar sua vida diária?	1	2	3	4	5
F11.4 Em que medida a sua qualidade de vida depende do uso de medicamentos ou de ajuda médica?	1	2	3	4	5
F13.1 Quão sozinho você se sente em sua vida?	1	2	3	4	5
F15.2 Quão satisfeitas estão as suas necessidades sexuais?	1	2	3	4	5
F15.4 Você se sente incomodado(a) por alguma dificuldade na sua vida sexual?	1	2	3	4	5
F16.1 Quão seguro(a) você se sente em sua vida diária?	1	2	3	4	5
F16.2 Você acha que vive em um ambiente seguro?	1	2	3	4	5
F16.3 O quanto você se preocupa com sua segurança?	1	2	3	4	5
F17.1 Quão confortável é o lugar onde você mora?	1	2	3	4	5
F17.4 O quanto você gosta de onde você mora?	1	2	3	4	5
F18.2 Você tem dificuldades financeiras?	1	2	3	4	5
F18.4 O quanto você se preocupa com dinheiro?	1	2	3	4	5
F19.1 Quão facilmente você tem acesso a bons cuidados médicos?	1	2	3	4	5
F21.3 O quanto você aproveita o seu tempo livre?	1	2	3	4	5
F22.1 Quão saudável é o seu ambiente físico (clima, barulho, poluição, atrativos)?	1	2	3	4	5
F22.2 Quão preocupado(a) você está com o barulho na área que você vive?	1	2	3	4	5

F23.2 Em que medida você tem problemas com transporte?	1	2	3	4	5
F23.4 O quanto às dificuldades de transporte dificultam sua vida?	1	2	3	4	5

nada	muito pouco	médio	muito	completamente
1	2	3	4	5

F2.1 Você tem energia suficiente para o seu dia-a-dia?	1	2	3	4	5
F7.1 Você é capaz de aceitar a sua aparência física?	1	2	3	4	5
F10.1 Em que medida você é capaz de desempenhar suas atividades diárias?	1	2	3	4	5
F11.1 Quão dependente você é de medicação?	1	2	3	4	5
F14.1 Você consegue dos outros o apoio que necessita?	1	2	3	4	5
F14.2 Em que medida você pode contar com amigos quando precisa deles?	1	2	3	4	5
F17.2 Em que medida as características de seu lar correspondem às suas necessidades?	1	2	3	4	5
F18.1 Você tem dinheiro suficiente para satisfazer suas necessidades?	1	2	3	4	5
F20.1 Quão disponível para você estão as informações que precisa no seu dia-a-dia?	1	2	3	4	5
F20.2 Em que medida você tem oportunidades de adquirir informações que considera necessárias?	1	2	3	4	5
F21.1 Em que medida você tem oportunidades de atividades de lazer?	1	2	3	4	5
F21.2 Quanto você é capaz de relaxar e curtir você mesmo?	1	2	3	4	5
F23.1 Em que medida você tem meios de transporte adequados?	1	2	3	4	5

muito insatisfeito	insatisfeito	nem satisfeito / nem insatisfeito	satisfeito	muito satisfeito
1	2	3	4	5

G2 Quão satisfeito(a) você está com a qualidade de sua vida?	1	2	3	4	5
G3 Em geral, quão satisfeito(a) você está com a sua vida?	1	2	3	4	5
G4 Quão satisfeito(a) você está com a sua saúde?	1	2	3	4	5
F2.3 Quão satisfeito(a) você está com a energia (disposição) que você tem?	1	2	3	4	5
F3.3 Quão satisfeito(a) você está com o seu sono?	1	2	3	4	5
F5.2 Quão satisfeito(a) você está com a sua capacidade de aprender novas informações?	1	2	3	4	5
F5.4 Quão satisfeito(a) você está com sua capacidade de tomar decisões?	1	2	3	4	5
F6.3 Quão satisfeito(a) você está consigo mesmo?	1	2	3	4	5
F6.4 Quão satisfeito(a) você está com suas capacidades?	1	2	3	4	5
F7.4 Quão satisfeito(a) você está com a aparência de seu corpo?	1	2	3	4	5
F10.3 Quão satisfeito(a) você está com sua capacidade de desempenhar as atividades do seu dia-a-dia?	1	2	3	4	5
F13.3 Quão satisfeito(a) você está com suas relações pessoais (amigos, parentes, conhecidos, colegas)?	1	2	3	4	5
F15.3 Quão satisfeito(a) você está com sua vida sexual?	1	2	3	4	5
F14.3 Quão satisfeito(a) você está com o apoio que você recebe de sua família?	1	2	3	4	5
F14.4 Quão satisfeito(a) você está com o apoio que você recebe de seus amigos?	1	2	3	4	5
F13.4 Quão satisfeito(a) você está com sua capacidade de dar apoio aos outros?	1	2	3	4	5
F16.4 Quão satisfeito(a) você está com a sua segurança física (assaltos, incêndios, etc.)?	1	2	3	4	5
F17.3 Quão satisfeito(a) você está com as condições do local onde mora?	1	2	3	4	5
F18.3 Quão satisfeito(a) você está com sua situação financeira?	1	2	3	4	5
F19.3 Quão satisfeito(a) você está com o seu acesso aos serviços de saúde?	1	2	3	4	5
F19.4 Quão satisfeito(a) você está com os serviços de assistência social?	1	2	3	4	5

F20.3 Quão satisfeito(a) você está com as suas oportunidades de adquirir novas habilidades?	1	2	3	4	5
F20.4 Quão satisfeito(a) você está com as suas oportunidades de obter novas informações?	1	2	3	4	5
F21.4 Quão satisfeito(a) você está com a maneira de usar o seu tempo livre?	1	2	3	4	5
F22.3 Quão satisfeito(a) você está com o seu ambiente físico (poluição, clima, barulho, atrativos)?	1	2	3	4	5
F22.4 Quão satisfeito(a) você está com o clima do lugar em que vive?	1	2	3	4	5
F23.3 Quão satisfeito(a) você está com o seu meio de transporte?	1	2	3	4	5
F13.2 Você se sente feliz com sua relação com as pessoas de sua família?	1	2	3	4	5

muito ruim	ruim	nem ruim / nem boa	boa	muito boa
1	2	3	4	5

G1 Como você avaliaria sua qualidade de vida?	1	2	3	4	5
F15.1 Como você avaliaria sua vida sexual?	1	2	3	4	5
F3.1 Como você avaliaria o seu sono?	1	2	3	4	5
F5.1 Como você avaliaria sua memória?	1	2	3	4	5
F19.2 Como você avaliaria a qualidade dos serviços de assistência social disponível para você?	1	2	3	4	5

Nunca	raramente	às vezes	repetidamente	sempre
1	2	3	4	5

F1.1 Com que frequência você sente dor (física)?	1	2	3	4	5
F4.2 Em geral, você se sente contente?	1	2	3	4	5
F8.1 Com que frequência você tem sentimentos negativos, tais como mau humor, desespero, ansiedade, depressão?	1	2	3	4	5

Nada	muito pouco	médio	muito	completamente
1	2	3	4	5

F12.1 Você é capaz de trabalhar?	1	2	3	4	5
F12.2 Você se sente capaz de fazer as suas tarefas?	1	2	3	4	5
F12.4 Quão satisfeito(a) você está com a sua capacidade para o trabalho?	1	2	3	4	5

muito ruim	ruim	nem ruim / nem boa	boa	muito boa
1	2	3	4	5

F12.3 Como você avaliaria a sua capacidade para o trabalho?	1	2	3	4	5
F9.1 Quão bem você é capaz de se locomover?	1	2	3	4	5

nada	muito pouco	mais ou menos	bastante	extremamente
1	2	3	4	5

F9.3 O quanto alguma dificuldade de locomoção lhe incomoda?	1	2	3	4	5
F9.4 Em que medida alguma dificuldade em mover-se afeta a sua vida no dia-a-dia?	1	2	3	4	5

Muito insatisfeito	insatisfeito	nem satisfeito nem insatisfeito	satisfeito	muito satisfeito
1	2	3	4	5

F9.2 Quão satisfeito(a) você está com sua capacidade de se locomover?	1	2	3	4	5
---	---	---	---	---	---

nada	muito pouco	mais ou menos	bastante	extremamente
1	2	3	4	5

F24.1 Suas crenças pessoais dão sentido à sua vida?	1	2	3	4	5
F24.2 Em que medida você acha que sua vida tem sentido?	1	2	3	4	5
F24.3 Em que medida suas crenças pessoais lhe dão força para enfrentar dificuldades?	1	2	3	4	5
F24.4 Em que medida suas crenças pessoais lhe ajudam a entender as dificuldades da vida?	1	2	3	4	5

**Anexo F – Instrumento de Avaliação de Qualidade de Vida desenvolvido pela
OMS – Versão abreviada (WHOQOL-BREVE)**

WHOQOL-BREVE

Por favor, leia cada questão, veja o que você acha e circule no número e lhe parece a melhor resposta.

	Muito ruim	Ruim	Nem ruim Nem boa	Boa	Muito boa
G1 Como você avaliaria sua qualidade de vida?	1	2	3	4	5

	Muito insatisfeito	Insatisfeito	Nem satisfeito Nem insatisfeito	Satisfeito	Muito satisfeito
G4 Quão satisfeito (a) você está com a sua saúde?	1	2	3	4	5

As questões seguintes são sobre **o quanto** você tem sentido algumas coisas nas últimas duas semanas.

	Nada	Muito pouco	Mais ou menos	Bastante	Extremamente
F1.4 Em que medida você acha que sua dor (física) impede você de fazer o que você precisa?	1	2	3	4	5
F11.3 Quanto você precisa de algum tratamento médico para levar sua vida diária?	1	2	3	4	5
F4.1 O quanto você aproveita a vida?	1	2	3	4	5
F24.2 Em que medida você acha que sua vida tem sentido?	1	2	3	4	5
F5.3 O quanto você consegue se concentrar?	1	2	3	4	5
F16.1 Quão seguro(a) você se sente em sua vida diária?	1	2	3	4	5
F22.1 Quão saudável é o seu ambiente físico (clima, barulho, poluição, atrativos) ?	1	2	3	4	5

As questões seguintes perguntam sobre **quão completamente** você tem sentido ou é capaz de fazer certas coisas nestas últimas duas semanas.

	Nada	Muito pouco	Médio	Muito	Completamente
F2.1 Você tem energia suficiente para o seu dia-a-dia?	1	2	3	4	5
F7.1 Você é capaz de aceitar a sua aparência física?	1	2	3	4	5
F18.1 Você tem dinheiro suficiente para satisfazer suas necessidades?	1	2	3	4	5
F20.1 Quão disponível para você estão às informações que precisa no seu dia-a-dia?	1	2	3	4	5
F21.1 Em que medida você tem oportunidades de atividades de lazer?	1	2	3	4	5

As questões seguintes perguntam sobre quão bem ou satisfeito você se sentiu a respeito de vários aspectos de sua vida nas últimas duas semanas.

	Muito ruim	Ruim	Nem ruim Nem bom	Bom	Muito bom
F9.1 Quão bem você é capaz de se locomover?	1	2	3	4	5

	Muito insatisfeito	Insatisfeito	Nem satisfeito Nem insatisfeito	Satisfeito	Muito satisfeito
F3.3 Quão satisfeito(a) você está com o seu sono?	1	2	3	4	5
F10.3 Quão satisfeito(a) você está com sua capacidade de desempenhar as atividades do seu dia-a-dia?	1	2	3	4	5
F12.4 Quão satisfeito(a) você está com a sua capacidade para o trabalho?	1	2	3	4	5
F6.3 Quão satisfeito(a) você está consigo mesmo?	1	2	3	4	5
F13.3 Quão satisfeito(a) você está com suas relações pessoais (amigos, parentes, conhecidos, colegas)?	1	2	3	4	5
F15.3 Quão satisfeito(a) você está com sua vida sexual?	1	2	3	4	5
F14.4 Quão satisfeito(a) você está com o apoio que você recebe de seus amigos?	1	2	3	4	5
F17.3 Quão satisfeito(a) você está com as condições do local onde mora?	1	2	3	4	5
F19.3 Quão satisfeito(a) você está com o seu acesso aos serviços de saúde?	1	2	3	4	5
F23.3 Quão satisfeito(a) você está com o seu meio de transporte?	1	2	3	4	5

As questões seguintes referem-se a com que frequência você sentiu ou experimentou certas coisas nas últimas duas semanas.

	Nunca	Algumas vezes	Freqüentemente	Muito freqüentemente	Sempre
F8.1 Com que frequência você tem sentimentos negativos, tais como mau humor, desespero, ansiedade, depressão?	1	2	3	4	5

Anexo G – Estimativas das habilidades (traços latentes) dos idosos

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens

idoso	escore_EQVF	escore_WB	escore_W100
001	1,358056	0,644518	0,402834
002	0,237812	0,359411	0,101711
003	0,189252	0,373761	-0,119105
004	-0,592149	-0,856915	-0,588986
005	1,057626	0,080733	-0,522695
006	1,790121	-0,359266	-0,121686
007	-0,639392	0,351941	0,634000
008	-0,106744	-0,270488	0,443445
009	-1,019128	-0,120845	0,070791
010	-0,907224	1,092107	0,565219
011	0,452198	-0,273958	-0,157910
012	1,790121	1,956695	2,698703
013	-1,304636	0,348043	-0,191086
014	1,790121	1,126057	1,874246
015	0,452198	1,412724	0,576098
016	0,669328	-0,496128	-0,361219
017	0,187483	1,956695	1,139718
018	-0,136479	-0,294932	0,051325
019	-0,093485	-0,223792	-0,004122
020	0,412788	-1,155508	-1,018464
021	-0,163337	-0,606442	-0,466091
022	0,112020	1,158282	1,615597
023	1,455176	-0,105565	0,472757
024	1,455176	-0,241375	-0,236819
025	-0,718622	-0,341690	-0,576866
026	1,137106	1,896932	1,177392
027	-0,430128	1,016938	0,433126
028	0,948490	-0,975097	-0,813777
029	1,057626	1,158282	1,989259
030	1,790121	1,956695	0,982839
031	-0,921890	-0,292719	-0,567027
032	-1,114388	-0,776638	-0,966666
033	0,215908	0,191171	0,080689
034	-0,690444	-1,08475	-0,983950
035	-0,438305	1,217032	1,221606
036	-2,410402	-1,696402	-1,283535
037	-0,218927	-0,835645	-0,945031
038	-0,613982	-0,17069	-0,231409
039	0,359869	0,376755	0,092264

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore EQVF	escore WB	escore W100
040	0,394749	-0,377851	-0,597162
041	0,445917	1,572214	0,752548
042	1,455176	-0,090037	0,277002
043	1,218377	1,512839	1,932690
044	0,087326	-0,335411	-0,764239
045	1,790121	0,219587	0,139994
046	-0,050246	-1,11984	-0,848543
047	-1,772578	-0,015445	-0,018793
048	-0,674458	-0,275497	-0,102919
049	-0,244535	0,104356	0,577143
050	-0,073811	-0,292688	-0,200759
051	-1,237259	0,044957	-0,200599
052	0,182452	-0,721210	-0,530597
053	1,790121	1,180781	1,727313
054	-0,064259	0,118711	0,309763
055	-0,843987	-0,686966	-0,453131
056	-1,628585	-0,461802	-0,482771
057	0,082067	-0,824956	-0,718605
058	0,291578	0,394569	-0,135599
059	-0,493873	-0,876701	-0,894079
060	0,774802	-0,100646	0,034311
061	-0,879804	-0,522636	-0,597722
062	-0,073032	0,584401	1,985391
063	1,455176	0,415501	0,474525
064	0,494493	0,975880	0,646137
065	1,057626	2,106429	0,856044
066	1,121921	-1,146658	-1,241149
067	1,790121	1,800097	2,010717
068	1,358056	-0,783257	-1,015749
069	-0,261588	-0,893571	-0,742874
070	-1,134146	-1,740261	-1,595673
071	0,219415	-0,74154	-0,700089
072	-1,281463	-1,774341	-1,737559
073	0,253907	-1,533149	-1,357572
074	-1,125386	-0,88813	-0,837066
075	0,700489	0,461517	0,271249
076	-1,345688	-3,024809	-2,164122
077	-0,454507	1,338547	1,896496
078	0,035216	0,823166	0,366926

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore_EQVF	escore_WB	escore_W100
079	-0,691204	-1,355239	-1,227237
080	1,016235	0,845942	0,711866
081	-0,112575	-0,943399	-0,736133
082	0,112025	1,158282	1,409124
083	0,236348	-0,383277	-0,299323
084	0,721372	0,026285	1,009925
085	0,141555	0,292185	0,194229
086	-0,812183	-1,223854	-1,053766
087	0,038212	-1,46714	-1,432544
088	0,698319	1,956695	2,381781
089	0,468859	-0,404902	0,011984
090	-0,907195	-0,107831	0,104409
091	-0,661769	0,947961	0,358305
092	-0,561077	-1,26877	-1,030524
093	1,790121	1,956695	1,927821
094	-1,543523	2,233412	0,436931
095	0,390713	-0,268496	-0,227144
096	1,790121	1,956695	2,753084
097	1,790121	0,219155	1,118537
098	-0,708095	-0,460778	-0,263515
099	-0,454831	-0,165456	0,038884
100	1,526645	1,015694	1,553234
101	0,186248	1,021108	0,894808
102	-1,240167	-0,34601	-0,068011
103	1,790121	1,572214	3,158085
104	-0,759765	-1,329101	-1,038027
105	-0,272563	0,216757	-0,188329
106	-1,940756	-1,819566	-1,570918
107	0,157714	-0,554307	-0,381533
108	-1,941593	-1,200740	-1,101541
109	1,057626	1,800097	2,171735
110	1,790121	0,746071	0,699152
111	1,790121	0,873237	0,691755
112	1,532152	1,233810	2,001693
113	0,082067	-0,704072	-0,882410
114	-0,079426	1,956695	1,740033
115	0,236645	0,978368	0,614072
116	-0,419360	-0,821314	-0,731767
117	-0,332875	0,204498	-0,133966

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore EQVF	escore WB	escore W100
118	-0,285348	0,754165	0,529514
119	-0,844217	-0,051309	-0,283442
120	-0,682060	-0,332219	-0,168444
121	0,518673	1,956695	1,697153
122	-0,205849	-0,286929	-0,298807
123	1,057626	-0,293929	-0,363957
124	-1,539777	-1,556983	-2,232754
125	0,112020	-0,556832	-0,522515
126	-0,115540	1,178772	0,407656
127	-0,088098	0,744570	0,152871
128	0,452198	-0,367892	-0,330708
129	0,772481	0,553793	1,465552
130	1,057626	-0,330824	-0,158669
131	0,555187	1,241122	1,143434
132	1,057626	1,956695	2,753084
133	0,452198	0,853331	1,182456
134	0,456397	-0,321724	-0,288068
135	1,790121	1,956695	1,871782
136	0,518673	1,800097	2,568963
137	-0,454831	-1,403162	-1,236308
138	-1,003861	-2,131299	-1,685359
139	0,669328	-0,883252	-0,584517
140	0,093430	0,456955	0,222052
141	-1,567622	-1,57248	-1,549867
142	1,532152	-0,056233	-0,411701
143	-1,074664	-0,133058	-0,490852
144	-0,488687	0,073973	0,224448
145	-1,058054	-0,056442	-0,064133
146	1,790121	1,016938	0,218159
147	1,790121	0,997878	0,957977
148	1,455176	-0,908615	-1,118292
149	0,518673	-0,669143	-0,490328
150	1,057626	1,972743	0,270413
151	0,115189	1,956695	1,869627
152	-1,195409	-0,242776	-0,531943
153	0,356014	-0,403327	-0,258176
154	-0,686035	1,194169	1,427545
155	-0,083881	-0,575211	-0,236585
156	0,142581	-0,766623	-0,627612

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore EQVF	escore WB	escore W100
157	1,790121	0,598975	0,813564
158	0,219415	-0,703927	-0,491670
159	0,741825	0,661344	0,144466
160	0,269956	-0,17246	0,120744
161	1,790121	0,046245	-0,427334
162	1,790121	0,473383	0,575516
163	0,669328	1,773445	0,814355
164	0,960137	0,017469	-0,556524
165	1,057626	0,178081	-0,022436
166	-0,290638	-0,022603	0,387785
167	1,790121	1,956695	1,979061
168	-0,280177	-1,189843	-0,762623
169	-0,887204	-0,337612	-0,367076
170	1,358056	2,079207	1,526309
171	0,518673	1,956695	0,872215
172	-0,682060	-0,205494	0,041368
173	-0,697900	-0,906696	-0,714162
174	0,221046	1,271324	0,734067
175	-0,073032	-1,296289	-1,086875
176	-0,056312	1,180781	0,954787
177	0,271648	0,249659	0,475948
178	-0,990234	-1,381447	-1,212073
179	0,112022	0,822302	0,520603
180	0,432113	-0,357732	-0,155378
181	0,040299	0,291552	0,068531
182	-0,070621	-0,586843	-0,549701
183	-0,516574	0,675486	0,252299
184	0,669328	1,956695	1,018321
185	-1,423412	-2,147572	-1,337692
186	-0,343072	-0,598077	-0,998633
187	0,235238	0,721479	0,439145
188	-0,359676	0,581657	0,436974
189	0,669328	1,126057	1,903223
190	1,455176	-1,226282	-0,977465
191	0,948490	-0,003081	-0,142106
192	1,790121	0,281452	-0,233012
193	0,215980	0,065810	0,273421
194	-0,231304	0,174590	0,238553
195	0,112020	0,584210	1,091314

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore_EQVF	escore_WB	escore_W100
196	0,047075	-1,047728	-0,676327
197	-1,217777	-1,394589	-1,556348
198	0,035216	-0,722636	-1,111799
199	0,490861	2,055381	1,773774
200	-0,279164	-0,934408	-1,251827
201	1,455176	0,420713	0,900047
202	-1,284057	-0,403149	-0,588115
203	-0,404988	0,169724	-0,556421
204	-0,403711	0,703472	0,569022
205	-1,214124	-0,874912	-0,939982
206	-1,298848	-0,820958	-1,112884
207	-0,870595	-0,503555	-0,693245
208	-0,107153	1,956695	0,938919
209	-0,560429	-0,872775	-1,004734
210	-0,223254	-0,713693	-0,633942
211	0,518673	0,321403	0,409665
212	0,069178	-0,803363	-0,608981
213	-0,098107	-0,181024	0,168301
214	-0,073032	1,217032	0,882285
215	1,790121	0,824236	0,483955
216	0,669328	0,768593	0,756662
217	-1,079736	-2,259019	-1,821867
218	-1,079736	-0,614963	-0,550636
219	-0,467529	0,676345	0,193749
220	-0,467529	-0,429242	-0,397274
221	0,401465	0,746301	0,428364
222	-0,540898	0,111225	-0,075706
223	-0,540898	0,859475	0,160901
224	-0,056315	-0,194878	-0,134883
225	1,790121	0,149963	0,705075
226	1,455176	1,355851	0,703483
227	-0,948753	0,103719	0,176418
228	0,093436	-0,679266	-0,365716
229	-0,603275	-0,353985	-0,135269
230	-0,797443	-1,64531	-1,499108
231	0,112024	-0,210489	0,145045
232	0,112028	-0,295731	-0,491728
233	-1,155751	-0,294693	-0,606279
234	-1,155751	-1,485371	-1,072751

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore_EQVF	escore_WB	escore_W100
235	-1,339043	-0,752305	-0,417461
236	1,790121	0,85036	1,124734
237	0,411134	0,85036	2,008464
238	1,790121	1,800097	1,804798
239	0,401465	0,513774	0,162949
240	-2,569301	-0,695659	-1,165305
241	-1,178824	-0,206525	-0,525187
242	-0,632116	-1,243065	-0,980565
243	-1,346435	-1,712075	-1,338486
244	-4,000000	-3,226655	-2,985033
245	0,669328	0,702587	0,595039
246	0,271648	0,886819	0,189414
247	-0,686817	-0,01296	-0,025785
248	-1,941593	-0,782102	-1,200921
249	-0,766498	-0,596574	-0,601485
250	-1,513165	-1,149566	-1,555126
251	-0,715162	-0,586419	-0,712357
252	-0,542378	0,174637	0,421694
253	0,307949	0,668492	0,289116
254	-2,410402	-2,750596	-2,161875
255	-4,000000	-2,972571	-2,377154
256	-0,944360	-0,705861	-0,585921
257	-0,540118	0,477834	0,203566
258	0,397464	-0,507473	-0,405692
259	1,455176	0,633159	2,091332
260	1,790121	0,076599	-0,068387
261	0,235238	-0,021188	0,137266
262	-0,467678	0,112776	-0,387087
263	-0,043164	1,956695	0,792496
264	1,790121	1,956695	1,874508
265	-1,415458	-1,652706	-1,238336
266	0,075181	0,071973	0,225802
267	-0,188503	0,541315	0,413954
268	-0,218927	0,104701	0,012414
269	1,455176	0,910183	1,098695
270	0,112020	1,956695	1,415509
271	-0,391971	-0,038112	0,038474
272	-0,294103	-0,489473	-0,447411
273	-0,502695	-0,388617	-0,338013

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore_EQVF	escore_WB	escore_W100
274	0,063874	-0,66347	-0,684983
275	-1,042408	-0,426708	-0,558104
276	-0,894266	-0,65412	-0,462913
277	-0,567210	-0,021415	0,461494
278	1,790121	0,125072	-0,501880
279	0,456397	1,191225	1,812606
280	-0,470749	-0,325945	-0,165980
281	1,790121	0,518461	1,151108
282	-0,306989	-0,605156	-0,329283
283	-0,439831	-0,435241	-0,192319
284	-1,049807	-0,736826	-0,347109
285	-1,150838	-0,110883	-0,092659
286	-1,050600	-0,342434	-0,365094
287	-2,617629	-1,223337	-1,473858
288	-0,596388	-2,113300	-1,692011
289	-1,592894	-1,713634	-1,882962
290	-0,823651	-1,299032	-1,399182
291	0,112020	-0,099431	-0,191196
292	-0,070247	-1,121855	-0,859313
293	-0,879486	-0,199166	-0,583101
294	-3,268569	-1,364932	-1,093445
295	-0,777826	-0,399313	-0,423949
296	-1,081077	-0,24352	-0,578160
297	-0,047152	-0,173473	0,104214
298	1,137106	1,572214	0,482982
299	1,790121	1,296001	0,902552
300	-0,540118	0,251046	1,559997
301	1,790121	1,709301	3,099734
302	1,790121	1,956695	1,649814
303	1,790121	1,234773	1,229555
304	-2,011489	-1,883841	-1,637949
305	-0,704333	-0,349364	-0,578877
306	1,057626	0,403273	0,628230
307	1,358056	0,829296	0,988901
308	1,790121	0,998303	0,377982
309	1,790121	2,233412	3,070026
310	1,790121	0,882497	0,803568
311	1,790121	1,956695	1,856208
312	1,790121	2,055381	2,148644

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore_EQVF	escore_WB	escore_W100
313	1,790121	1,956695	2,357432
314	-1,395224	-1,059055	-0,835986
315	1,358056	0,533094	1,097201
316	1,790121	1,047766	1,135019
317	1,790121	0,903928	0,438197
318	1,057626	0,600066	0,986127
319	0,104564	1,300228	0,648224
320	-1,064501	-1,208411	-1,004403
321	1,790121	0,078975	0,053105
322	-1,215148	0,350731	0,322208
323	1,057626	0,841853	0,319214
324	-0,360417	-0,412683	-0,552121
325	0,600321	0,643367	0,598518
326	1,526645	2,055381	2,534074
327	-0,213917	0,454859	0,247815
328	1,790121	0,786366	0,308223
329	1,218377	1,956695	2,753084
330	0,518673	0,219556	0,026715
331	-0,136479	-0,291095	0,170654
332	1,532152	0,746071	0,418412
333	1,790121	0,641449	0,733323
334	1,057626	1,213176	0,718604
335	1,016235	0,785047	0,912228
336	0,109423	0,469553	0,661039
337	1,790121	0,905571	1,415425
338	0,155698	0,159944	0,215017
339	-1,362605	-1,227599	-1,234444
340	-4,000000	-4,000000	-4,000000
341	-0,462432	-0,255454	0,493537
342	-1,346430	-2,517644	-2,398747
343	-0,307531	-0,599211	-0,716770
344	-0,686878	-0,884616	-0,790351
345	-0,922968	-1,456669	-0,975724
346	-1,245833	-0,788425	-1,176709
347	-1,749638	-1,724821	-1,645125
348	-0,057468	-0,185145	-0,167709
349	-1,013172	-0,308832	-0,746247
350	-0,213917	0,062723	0,088648
351	-0,807693	-0,352967	-0,264036

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore_EQVF	escore_WB	escore_W100
352	-2,410402	-0,56288	-0,538811
353	-1,885011	-0,437061	-0,684514
354	-0,045096	-0,377906	-0,522461
355	-1,161632	-1,04977	-0,997641
356	-1,079736	-0,467169	-0,137766
357	0,215985	-0,382546	-0,218457
358	0,456397	0,583556	0,448452
359	0,465457	0,536931	0,307217
360	0,419672	2,055381	1,558343
361	-1,543953	-2,003412	-1,391656
362	0,892825	-0,411526	-0,540151
363	-1,403806	0,357697	-0,073368
364	-0,425314	0,252549	0,544358
365	-2,292737	-2,240179	-2,380674
366	-3,268569	-3,394047	-2,647484
367	0,250374	-1,008721	-0,589122
368	-1,470252	-1,635864	-1,486689
369	-1,263259	-1,392756	-1,246644
370	-1,139426	-1,103349	-0,900419
371	-0,529572	-0,270399	-0,436531
372	0,518673	-0,270399	-0,488891
373	-0,620999	-0,202378	-0,010386
374	0,607321	-0,145456	-0,003178
375	-0,293100	-0,117078	-0,504742
376	-2,022844	-2,037923	-1,919717
377	0,494435	1,197156	0,326632
378	1,455176	-0,320542	-0,165779
379	1,790121	0,808835	1,797672
380	-0,001768	1,956695	2,149082
381	0,419672	0,001918	0,121279
382	1,790121	1,290091	1,691141
383	-0,498018	-1,799099	-1,454632
384	0,700481	-0,232251	-0,570549
385	1,790121	1,956695	2,123974
386	1,790121	1,956695	0,856292
387	-0,244531	-0,713399	-0,565984
388	-1,334585	-2,189953	-1,635432
389	1,790121	0,487753	1,918863
390	0,492934	0,036841	0,298937

TABELA 1 – Estimativas das habilidades (traços latentes globais) dos instrumentos EQVF, WHOQOL-BREVE e WHOQOL-100, respectivamente com 14, 23 e 65 itens (continuação)

idoso	escore_EQVF	escore_WB	escore_W100
391	-1,607985	-2,013944	-2,092978
392	-0,728605	1,178772	0,711904
393	-2,139662	-1,234141	-0,852071
394	-0,920241	-0,327975	-0,301159
395	-0,620611	-0,137647	-0,713277
396	0,112028	0,423373	0,750243
397	0,215988	-0,602161	-0,484524
398	-0,347223	0,666016	-0,185889
399	-1,289633	-0,606679	-0,756844
400	-1,248743	1,291795	0,132608

Anexo H – Tabelas que não foram colocadas no artigo

TABELA 2 – Análise Fatorial para a Escala de Qualidade de Vida de Flanagan (EQVF), sem item 5.

Item	Fator 1	Fator 2	Fator 3
09- Aprendizagem	0,729	-0,027	0,023
08- Participação em associações	0,725	0,040	-0,007
13- Recreação ativa	0,677	0,094	0,140
12- Comunicação	0,555	0,348	0,200
10- Auto-conhecimento	0,522	0,198	0,279
07- Voluntariado	0,474	0,301	0,139
11- Trabalho	0,460	0,315	0,292
14- Recreação passiva	0,345	0,337	0,141
03- Relacionamento familiar	0,034	0,724	0,101
06- Amizade	0,217	0,678	0,091
15- Socialização	0,350	0,593	-0,120
04- Construir família	-0,018	0,577	0,163
02- Saúde	0,228	0,023	0,787
01- Conforto material	0,072	0,184	0,719
Autovalor	4,138	1,318	1,059
% da variação total	29,556	9,411	7,561

KMO: 0,859

Teste de Bartlett : $p < 0,0001$

Método de extração: Análise de componentes principais. Método de Rotação: Varimax com normalização de Kaiser.

TABELA 3 - Análise Fatorial para WHOQOL – Breve, sem os itens F153, F181, F193.

Item	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	Fator 6
F21 (energia para dia-a-dia)	0,714	0,224	0,135	0,104	-0,009	-0,107
F103 (atividades do dia-a-dia)	0,693	0,044	0,128	0,195	0,192	0,035
G4 (satisfação com saúde)	0,651	0,154	0,069	0,113	0,028	0,266
F113 (tratamento médico)	0,629	0,124	-0,200	-0,001	0,054	0,351
F91 (locomoção)	0,575	0,064	0,122	-0,047	0,381	0,094
F124 (capacidade pro trabalho)	0,565	0,153	0,215	0,140	0,295	-0,021
F14 (capacidade impedida pela dor)	0,458	0,235	0,298	-0,121	0,061	0,231
F71 (aparência física)	0,412	-0,160	0,384	0,248	-0,125	-0,061
F242 (vida tem sentido)	0,328	0,207	0,210	0,204	0,250	-0,139
F211 (oportunidade de lazer)	0,203	0,758	0,025	0,113	0,052	-0,019
F201 (disponibilidade das informações)	0,086	0,607	0,007	0,094	0,166	0,007
F41 (aproveitar a vida)	0,160	0,510	0,391	0,109	0,093	0,227
F161 (segurança)	0,103	-0,029	0,707	0,074	0,268	0,111
F53 (concentração)	0,180	0,347	0,528	0,095	-0,308	0,089
F81 (sentimentos negativos)	0,081	0,327	0,409	0,166	0,176	0,355
F144 (apoio dos amigos)	0,119	0,194	-0,035	0,789	0,017	0,056
F133 (relações pessoais)	0,044	0,092	0,185	0,780	0,080	-0,039
F63 (satisfação consigo mesmo)	0,362	0,039	0,332	0,453	0,182	0,148
F173 (satisfação com local de morada)	0,032	0,193	0,001	0,158	0,669	0,247
F233 (satisfação com transporte)	0,231	0,052	0,114	-0,020	0,537	-0,076
F221 (ambiente físico)	0,171	0,317	0,350	0,100	0,356	-0,260
G1 (qualidade de vida)	0,240	0,296	-0,015	0,279	0,353	0,294
F33 (satisfação com sono)	0,161	-0,011	0,163	0,011	0,034	0,756
Autovalores	5,930	1,534	1,227	1,150	1,042	1,031
% da variação total	25,784	6,667	5,336	4,989	4,530	4,482

KMO: 0,896

Teste de Bartlett: $p < 0,0001$

Método de extração: Análise de componentes principais. Método de Rotação: Varimax com normalização de Kaiser.

TABELA 4 – Análise Fatorial para WHOQOL-100 sem os 35 itens inconsistentes.

Faceta	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5
Faceta 12	0,815	0,183	0,088	0,062	0,041
Faceta 10 (sem itens F102 e F104)	0,792	0,139	0,105	0,047	0,219
Faceta 2 (sem itens F22 e F24)	0,734	0,184	0,117	0,199	0,105
Faceta 9 (sem itens F93 e F94)	0,696	-0,020	0,301	0,103	0,048
Faceta 6	0,539	0,440	0,117	0,206	0,146
Faceta Geral	0,537	0,393	0,347	0,240	-0,012
Faceta 24	0,456	0,373	-0,113	0,162	-0,054
Faceta 14	0,006	0,764	0,179	0,110	0,083
Faceta 13 (sem item F131)	0,245	0,747	0,091	-0,007	0,116
Faceta 21	0,303	0,549	0,296	0,320	-0,054
Faceta 4	0,476	0,546	0,191	0,311	-0,004
Faceta 17	0,202	0,485	0,404	-0,117	0,188
Faceta 19 (sem item F191)	0,125	0,192	0,660	0,056	-0,036
Faceta 18 (sem itens F181, F182 e F184)	-0,050	0,151	0,648	0,055	-0,001
Faceta 23 (sem itens F232 e F234)	0,305	-0,026	0,616	0,063	0,057
Faceta 16 (sem itens F163 e F164)	0,136	0,131	0,478	0,116	0,434
Faceta 20	0,201	0,231	0,455	0,398	0,074
Faceta 3 (sem itens F32 e F34)	0,170	-0,017	0,114	0,637	0,013
Faceta 15 (sem itens F152, F153 e F154)	0,041	0,150	-0,110	0,575	0,359
Faceta 5	0,381	0,218	0,287	0,495	-0,014
Faceta 7 (sem itens F72 e F73)	0,092	0,037	-0,016	0,157	0,809
Faceta 22 (sem item F222)	0,269	0,344	0,356	-0,202	0,423
Autovalores	7,256	1,581	1,242	1,128	1,060
% da variação total	32,980	7,184	5,645	5,126	4,818

KMO: 0,915

Teste de Bartlett: $p < 0,0001$

Método de extração: Análise de componentes principais. Método de Rotação: Varimax com normalização de Kaiser.

TABELA 5 – Correlações entre os escores latentes (E.L) globais (G) e dos domínios (D) entre os instrumentos WHOQOL-BREVE (dos itens inconsistentes).

	WB(G26)	WB(G23)	WB(D1)	WB(D2)	WB(D3)	WB(D4)	W100(G65)	W100(D1)	W100(D2)	W100(D3)	W100(D4)	W100(D5)	W100(D6)
WB(G26)	1	0,975**	0,814**	0,748**	0,502**	0,405**	0,911**	-0,564**	0,783**	0,783**	0,783**	0,783**	0,783**
WB(G23)		1	0,854**	0,787**	0,488**	0,422**	0,916**	-0,595**	0,805**	0,805**	0,805**	0,805**	0,805**
WB(D1)			1	0,554**	0,351**	0,271**	0,758**	-0,623**	0,649**	0,649**	0,649**	0,649**	0,649**
WB(D2)				1	0,399**	0,361**	0,720**	-0,446**	0,785**	0,785**	0,785**	0,785**	0,785**
WB(D3)					1	0,712**	0,490**	-0,141**	0,423**	0,423**	0,423**	0,423**	0,423**
WB(D4)						1	0,415**	-0,104*	0,361**	0,361**	0,361**	0,361**	0,361**
W100(G65)							1	-0,511**	0,845**	0,845**	0,845**	0,845**	0,845**
W100(D1)								1	-0,464**	-0,464**	-0,464**	-0,464**	-0,464**
W100(D2)									1	1	1	1	1
W100(D3)													
W100(D4)													
W100(D5)													
W100(D6)													

** Significativo a 1%

* Significativo a 5%

Anexo I – Resultados adicionais que não foram colocados no artigo

Análises de Variância:

EQVF:

Dependent Variable: ESCORE BRUTO 14 ITENS

> F	Source	DF	Sum of Squares	Mean Square	F Value	Pr
<.0001	Model	11	1435.707031	130.5189119	17.49	
0.0846	SEX0	1	22.2892710	22.2892710	2.99	
<.0001	DEP_CAT	2	1276.626530	638.313265	85.49	
0.0011	DESESPER	3	122.500236	40.833412	5.47	
0.7840	IDADE_K	2	3.635998	1.817999	0.24	
0.9197	ESCOL_K	1	0.075920	0.075920	0.01	
0.2346	EST_CIV2	1	10.580076	10.580076	1.42	
0.0053	SAUDE_K	1	58.592597	58.592597	7.85	
	Error	388	2896.899701	7.466236		
	Corrected Total	399	4332.606732			

R-Square	Coeff Var	Root MSE	BR_SEM5 Mean
0.347354	28.30812	2.732441	9.652500

SEX0	BR_SEM5 LSMEAN
1.00	6.22531046
2.00	6.43716542

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DEP_CAT	BR_SEM5 LSMEAN	LSMEAN Number
1.00	9.06726433	1
2.00	6.42025313	2
3.00	4.25619636	3

Least Squares Means for effect DEP_CAT
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: BR_SEM5			
i/j	1	2	3
1		<.0001	<.0001
2	<.0001		0.0168
3	<.0001	0.0168	

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DESESPER	BR_SEM5 LSMEAN	LSMEAN Number
1.00	7.97625673	1
2.00	7.05334922	2
3.00	6.76526393	3
4.00	4.53008189	4

Least Squares Means for effect DESESPER
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: BR_SEM5

i / j	1	2	3	4
1		0.0517	0.2475	0.0170
2	0.0517		0.9725	0.1350
3	0.2475	0.9725		0.2541
4	0.0170	0.1350	0.2541	

Least Squares Means

I DADE_K	BR_SEM5 LSMEAN	LSMEAN Number
1.00	6.78260736	1
2.00	6.74052312	2
3.00	6.22058335	3

Least Squares Means

ESCOL_K	BR_SEM5 LSMEAN
1.00	6.67434840
2.00	6.48812749

Least Squares Means

EST_CIV2	BR_SEM5 LSMEAN
1.00	6.41042175
2.00	6.75205413

Least Squares Means

SAUDE_K	BR_SEM5 LSMEAN
1.00	6.17330456
2.00	6.98917132

Dependent Variable: ESCORE LATENTE SEM ITEM 5 (14 ITENS)

> F	Source	DF	Sum of Squares	Mean Square	F Value	Pr
<.0001	Model	11	139.5937899	13.2357991	14.01	
0.0614	SEX0	1	3.3212084	3.3212084	3.52	
<.0001	DEP_CAT	2	119.3641084	59.6820542	63.19	
0.1327	DESESPER	3	6.4323755	2.1441252	2.27	
0.3308	I DADE_K	2	2.0956475	1.0478237	1.11	
0.8054	ESCOL_K	1	0.0574047	0.0574047	0.06	
0.1389	EST_CIV2	1	2.0768389	2.0768389	2.20	
0.0100	SAUDE_K	1	6.2462065	6.2462065	6.61	

Error	388	366.4393581	0.9444313
Corrected Total	399	506.0331479	

R-Square	Coeff Var	Root MSE	ESC_SEM5 Mean
0.284344	17227.34	0.971819	0.005641

Least Squares Means
ESC_SEM5
LSMEAN

SEXO	
1.00	-1.03648748
2.00	-0.83117706

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DEP_CAT	ESC_SEM5 LSMEAN	LSMEAN Number
1.00	-0.15803918	1
2.00	-0.96946427	2
3.00	-1.67399337	3

Least Squares Means for effect DEP_CAT
Pr > |t| for H0: LSMEAN(i)=LSMEAN(j)

Dependent Variable: ESC_SEM5

i / j	1	2	3
1		<.0001	<.0001
2	<.0001		0.0321
3	<.0001	0.0321	

Least Squares Means
ESC_SEM5
LSMEAN

DESESPER		LSMEAN Number
1.00	-0.57309627	1
2.00	-0.85805116	2
3.00	-0.90967705	3
4.00	-1.39450461	4

Least Squares Means
ESC_SEM5
LSMEAN

IDADE_K		LSMEAN Number
1.00	-0.79117855	1
2.00	-0.83964374	2
3.00	-1.17067453	3

Least Squares Means
ESC_SEM5
LSMEAN

ESCOL_K	
1.00	-0.92464594
2.00	-0.94301861

Least Squares Means
ESC_SEM5
LSMEAN

EST_CIV2	
1.00	-1.01003973
2.00	-0.85762482

Least Squares Means

SAUDE_K	ESC_SEM5 LSMEAN
1.00	-1.06702363
2.00	-0.80064092

WBREVE:

Dependent Variable: ESCORE BRUTO COM 23 ITENS

> F	Source	DF	Sum of Squares	Mean Square	F Value	Pr
<.0001	Model	11	4792.46218	435.67838	28.24	
0.9764	SEXO	1	0.013567	0.013567	0.00	
0.8045	IDADE_K	2	6.717770	3.358885	0.22	
0.0524	ESCOL_K	1	58.398337	58.398337	3.79	
0.8875	EST_CIVK	1	0.309433	0.309433	0.02	
<.0001	SAUDE_K	1	1419.649325	1419.649325	92.01	
<.0001	DEP_CAT	2	2916.589628	1458.294814	94.52	
<.0001	DESESPER	3	390.784123	130.261374	8.44	
	Error	388	5986.32782	15.42868		
	Corrected Total	399	10778.79000			

R-Square	Coeff Var	Root MSE	SOMA23IT Mean
0.444620	25.02667	3.927936	15.69500

Least Squares Means	
SEXO	SOMA23IT LSMEAN
1.00	11.4889392
2.00	11.1738499

Least Squares Means	
IDADE_K	SOMA23IT LSMEAN
1.00	11.0760155
2.00	11.3757193
3.00	11.5424488

Least Squares Means	
ESCOL_K	SOMA23IT LSMEAN
1.00	11.4828468
2.00	11.1799422

Least Squares Means	
EST_CIVK	SOMA23IT LSMEAN
1.00	11.1652413
2.00	11.4975478

Least Squares Means	
SAUDE_K	SOMA23IT LSMEAN

1.00	10.1264482
2.00	12.5363409

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DEP_CAT	SOMA23IT LSMEAN	LSMEAN Number
1.00	15.3814066	1
2.00	11.4720902	2
3.00	7.1406868	3

Least Squares Means for effect DEP_CAT
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: SOMA23IT

i / j	1	2	3
1		<.0001	<.0001
2	<.0001		0.0004
3	<.0001	0.0004	

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DESESPER	SOMA23IT LSMEAN	LSMEAN Number
1.00	13.8250577	1
2.00	11.8352286	2
3.00	10.4999761	3
4.00	9.1653157	4

Least Squares Means for effect DESESPER
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: SOMA23IT

i / j	1	2	3	4
1		0.0008	0.0024	0.0285
2	0.0008		0.4998	0.3836
3	0.0024	0.4998		0.8696
4	0.0285	0.3836	0.8696	

Dependent Variable: ESCORE LATENTE COM 23 ITENS

Source	DF	Sum of Squares	Mean Square	F Value	Pr
> F Model	11	197.7096307	17.9736028	23.67	<.0001
SEX0	1	0.0659841	0.0659841	0.09	0.7683
IDADE_K	2	1.4487981	0.7243990	0.95	0.3861
ESCOL_K	1	3.1969531	3.1969531	4.21	0.0508
EST_CIVK	1	0.0105439	0.0105439	0.01	0.9063
SAUDE_K	1	70.1868136	70.1868136	92.44	<.0001
DEP_CAT	2	110.1952526	55.0976263	72.57	<.0001

0.0010	DESESPER	3	12.6052854	4.2017618	5.53
	Error	388	294.5898008	0.7592521	
	Corrected Total	399	492.2994316		

R-Square	Coeff Var	Root MSE	ESC_231 T Mean
0.401604	5264.809	0.871351	0.016550

Least Squares Means
ESC_231 T
LSMEAN

SEXO	
1.00	-0.77572841
2.00	-0.86237043

Least Squares Means

IDADE_K	ESC_231 T LSMEAN	LSMEAN Number
1.00	-0.81597290	1
2.00	-0.82524745	2
3.00	-0.81592791	3

Least Squares Means
ESC_231 T
LSMEAN

ESCOL_K	
1.00	-0.79438331
2.00	-0.84371553

Least Squares Means
ESC_231 T
LSMEAN

EST_CIVK	
1.00	-0.85962904
2.00	-0.77846981

Least Squares Means

SAUDE_K	ESC_231 T LSMEAN
1.00	-1.10983982
2.00	-0.52825902

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DEP_CAT	ESC_231 T LSMEAN	LSMEAN Number
1.00	-0.01679846	1
2.00	-0.80531039	2
3.00	-1.63503942	3

Least Squares Means for effect DEP_CAT
Pr > |t| for H0: LSMEAN(i)=LSMEAN(j)

Dependent Variable: ESC_231 T

i / j	1	2	3
1		<.0001	<.0001
2	<.0001		0.0029
3	<.0001	0.0029	

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DESESPER	ESC_231 T LSMEAN	LSMEAN Number
1.00	-0.37818698	1
2.00	-0.75385938	2
3.00	-0.93086156	3
4.00	-1.21328977	4

Least Squares Means for effect DESESPER
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: ESC_231 T

i / j	1	2	3	4
1		0.0063	0.0403	0.1116
2	0.0063		0.8369	0.6043
3	0.0403	0.8369		0.8846
4	0.1116	0.6043	0.8846	

WHOQOL-100:

Dependent Variable: ESCORE BRUTO COM 65 ITENS

> F	Source	DF	Sum of Squares	Mean Square	F Value	Pr
<.0001	Model	11	29139.33365	2649.03033	24.90	
0.3376	SEXO	1	98.08068	98.08068	0.92	
<.0001	DEP_CAT	2	22582.73245	11291.36622	106.14	
<.0001	DESESPER	3	4067.22571	1355.74190	12.74	
0.8094	IDADE_K	2	45.01384	22.50692	0.21	
0.8898	ESCOL_K	1	2.04337	2.04337	0.02	
<.0001	SAUDE_K	1	2317.87932	2317.87932	21.79	
0.6189	EST_CIVK	1	26.35828	26.35828	0.25	
	Error	388	41275.90635	106.38120		
	Corrected Total	399	70415.24000			

R-Square	Coeff Var	Root MSE	ESC_651 T Mean
0.413821	23.11030	10.31413	44.63000

Least Squares Means
ESC_651 T
LSMEAN

SEXO	LSMEAN
1.00	32.5208897
2.00	32.8955837

ESC_651 T LSMEAN

DEP_CAT	LSMEAN	Number
1.00	42.3962850	1
2.00	33.0860610	2
3.00	22.6423642	3

Least Squares Means for effect DEP_CAT
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: ESC_65I T

i / j	1	2	3
1		<.0001	<.0001
2	<.0001		0.0014
3	<.0001	0.0014	

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DESESPER	ESC_65I T LSMEAN	LSMEAN Number
1.00	40.0226924	1
2.00	34.2062356	2
3.00	31.1587728	3
4.00	25.4452461	4

Least Squares Means for effect DESESPER
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: ESC_65I T

i / j	1	2	3	4
1		0.0001	0.0020	0.0054
2	0.0001		0.6162	0.1931
3	0.0020	0.6162		0.5959
4	0.0054	0.1931	0.5959	

IDADE_K	Least Squares Means LSMEAN	Number
1.00	32.2140129	1
2.00	32.8177927	2
3.00	33.0929045	3

ESCOL_K	Least Squares Means ESC_65I T LSMEAN
1.00	33.3570319
2.00	32.0594416

SAUDE_K	Least Squares Means ESC_65I T LSMEAN
1.00	30.1456426
2.00	35.2708309

EST_CIVK	Least Squares Means ESC_65I T LSMEAN
1.00	32.4314743
2.00	32.9849992

Dependent Variable: ESCORE LATENTE COM 65 ITENS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	154.3630228	14.0330021	17.17	<.0001
SEXO	1	0.1471884	0.1471884	0.18	0.6715
DEP_CAT	2	113.9314398	56.9657199	69.70	<.0001
DESESPER	3	22.7548041	7.5849347	9.28	<.0001
IDADE_K	2	0.5806170	0.2903085	0.36	0.7013
ESCOL_K	1	0.0033745	0.0033745	0.00	0.9488
SAUDE_K	1	16.5144959	16.5144959	20.20	<.0001
EST_CIVK	1	0.4311031	0.4311031	0.53	0.4681
Error	388	317.1309885	0.8173479		
Corrected Total	399	471.4940113			

R-Square	Coeff Var	Root MSE	LATEN_65 Mean
0.327391	3118.917	0.904073	0.028987

Least Squares Means
LATEN_65
LSMEAN

SEXO	LSMEAN
1.00	-0.78722674
2.00	-0.80956977

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DEP_CAT	LATEN_65 LSMEAN	LSMEAN Number
1.00	-0.14305345	1
2.00	-0.81911562	2
3.00	-1.43302568	3

Least Squares Means for effect DEP_CAT
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: LATEN_65

i / j	1	2	3
1		<.0001	<.0001
2	<.0001		0.0486
3	<.0001	0.0486	

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

DESESPER	LATEN_65 LSMEAN	LSMEAN Number
1.00	-0.28852898	1
2.00	-0.72299892	2
3.00	-0.94938019	3
4.00	-1.23268492	4

Least Squares Means for effect DESESPER
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: LATEN_65				
i / j	1	2	3	4
1		0.0017	0.0124	0.0692
2	0.0017		0.7310	0.5497
3	0.0124	0.7310		0.8944
4	0.0692	0.5497	0.8944	

Least Squares Means

IDADE_K	LATEN_65 LSMEAN	LSMEAN Number
1.00	-0.74791360	1
2.00	-0.82557497	2
3.00	-0.82170619	3

Least Squares Means

ESCOL_K	LATEN_65 LSMEAN
1.00	-0.74852553
2.00	-0.84827098

Least Squares Means

SAUDE_K	LATEN_65 LSMEAN
1.00	-1.01457325
2.00	-0.58222325

Least Squares Means

EST_CIVK	LATEN_65 LSMEAN
1.00	-0.83379303
2.00	-0.76300347

Anexo J – Estatísticas de concordância

EQVF:

BRUTO	LATENTE		Total
Frequency,	0,	1,	
0	200	13	213
1	0	187	187
Total	200	200	400

Statistics for Table of BRUTO by LATENTE

McNemar's Test
 Statistic (S) 13.0000
 DF 1
 Pr > S 0.0003

Simple Kappa Coefficient
 Kappa 0.9350
 ASE 0.0177
 95% Lower Conf Limit 0.9003
 95% Upper Conf Limit 0.9697

Sample Size = 400

WBREVE:

BRUTO	LATENTE		Total
Frequency,	0,	1,	
0	193	9	202
1	7	191	198
Total	200	200	400

McNemar's Test
 Statistic (S) 0.2500
 DF 1
 Pr > S 0.6171

Simple Kappa Coefficient
 Kappa 0.9200
 ASE 0.0196
 95% Lower Conf Limit 0.8816
 95% Upper Conf Limit 0.9584

Sample Size = 400

W100:

Table of BRUTO by LATENTE

BRUTO	LATENTE		Total
Frequency,	0,	1,	
0	189	11	200
1	12	188	200

Total	201	199	400
-------	-----	-----	-----

Statistics for Table of BRUTO by LATENTE

McNemar's Test
 ffffffffffffffffffffffffff
 Statistic (S) 0.0435
 DF 1
 Pr > S 0.8348

Simple Kappa Coefficient
 ffffffffffffffffffffffffff
 Kappa 0.8850
 ASE 0.0233
 95% Lower Conf Limit 0.8394
 95% Upper Conf Limit 0.9306

Sample Size = 400

Anexo K – Recursos Computacionais

RECURSOS COMPUTACIONAIS

O crescimento e a divulgação da TRI estão extremamente ligados ao desenvolvimento de ferramentas computacionais específicas, viabilizando sua aplicação. A TRI é uma técnica de análise complexa e exige procedimentos dispendiosos, ferramentas matemáticas muito mais elaboradas que as empregadas na TCM ou em outras técnicas já conhecidas.

Em TRI, desde suas primeiras aplicações, pesquisadores desenvolveram softwares específicos para seus problemas, mas a difusão da técnica depende diretamente da disponibilização desses programas computacionais. Na Europa e nos EUA, onde a TRI já vem sendo usada a mais tempo, existe um maior número de programas computacionais disponíveis desde a década de 70, mas no Brasil, onde usa-se essa ferramenta a pouco mais de 15 anos, ainda são poucos conhecidos e usados tais softwares. Existem programas para análise de dados dicotômicos (BILOG, BILOG-MG, TESTFACT, ...) e para dados politômicos (PARSCALE, MULTILOG, RUMM, ...). Nessa seção serão abordados os programas utilizados nesse trabalho, PARSCALE versão 4.1 e BILOG.

1. PARSCALE

O PARSCALE foi desenvolvido nos anos 90 por Eiji Muraki, do Educational Testing Service dos Estados Unidos, e R. Darrel Bock, da Universidade de Chicago, e nele foram implementados métodos para análise de respostas categóricas, sejam elas binárias ou politômicas.

O programa permite lidar com um amplo espectro de tipos de dados, podendo inclusive ajustar modelos para blocos de itens, onde cada bloco segue uma estrutura diferente. Basicamente, ajusta os modelos de 1, 2 e 3 parâmetros, os modelos de resposta gradual, de resposta gradual modificado, de crédito parcial e de crédito parcial generalizado. O PARSCALE não ajusta o modelo de resposta nominal, mas apresenta características importantes que o outro aplicativo, o MULTILOG, não contempla, como por exemplo, para qualquer modelo, permite ao pesquisador computar a informação do item em vários pontos da escala e salvar esses valores em um arquivo texto externo, para futuras manipulações, além de estar incluído no mesmo um poderoso teste quiquadrado para ajuste do modelo.

Em termos de estimação de parâmetros, o usuário pode escolher se deseja que os parâmetros sejam estimados na métrica normal (default) ou logística. O usuário também pode escolher outras opções para detalhes de estimação dos parâmetros, como por exemplo, especificar o número de pontos de quadratura e seus pesos, mudar as distribuições a priori dos interceptos, inclinações e thresholds, e mudar o

número de ciclos de estimação. Também estão inclusos no programa os métodos de máxima verossimilhança, máxima verossimilhança ponderada e o método de esperança a posteriori.

O programa também apresenta a análise de comportamento diferencial (*Differential Item Functioning*), análises para tendências entre os grupos e ao longo do tempo, dentre outras funções importantes. O PARSCALE permite manipulações de arquivos de programação e uma revisão eficiente dos arquivos de saída. As análises são realizadas em 4 etapas, chamadas fase 0, 1, 2 e 3. Na fase 1, que é a entrada e leitura de dados, o usuário deve fornecer basicamente duas informações: a identificação do indivíduo e as respostas. O programa lê apenas arquivos em formato ASCII (.dat), que são arquivos de texto, onde primeiro identificam-se os indivíduos, seguido das respostas dos itens sem separação de caracteres. Na fase 0, são emitidas as informações dos dados, número de itens, formato do banco de dados, especificações do modelo, parâmetros de calibração, pontos de quadratura e informações sobre os arquivos temporários e de saída. Também são mostradas as informações sobre os dois primeiros indivíduos, a fim de o usuário conferir se o programa está lendo adequadamente as informações. Na fase 1, são apresentadas as estatísticas descritivas para cada item, como número e porcentagem para cada categoria, resposta média e desvio padrão para cada item, correlação polisserial e as estimativas iniciais para os parâmetros. Essa fase é de suma importância, pois são fornecidos os valores que servem de diagnóstico preliminar dos itens.

A fase 2 é a fase de calibração dos itens, isto é, onde são estimados os parâmetros dos itens e seus respectivos erros padrões, além do teste quiquadrado de ajuste. E na fase 3 são estimados os escores do traço latente para cada sujeito a partir das estimativas iniciais da fase anterior. Esses traços latentes são estimados na mesma escala dos parâmetros, mas podem-se especificar algumas mudanças na escala, caso seja de interesse. Na janela de saída, são apresentados os escores para cada sujeito, valores para as CCI's e funções de informação, que serão usados na confecção dos gráficos das habilidades.

2. BILOG

O BILOG é um software específico para análise de dados binários e foi criado inicialmente para análises educacionais, e seu modo de apresentação das informações, portanto, segue tal estilo. Seu mecanismo é muito semelhante ao PARSCALE. Basicamente ajusta os modelos de 1, 2 e 3 parâmetros.

Na fase 1, de entrada e leitura dos dados, o usuário deve fornecer a identificação de cada indivíduo com suas respectivas respostas e o gabarito, que é uma seqüência contando as alternativas corretas dos itens que compõem o teste. Também é possível fornecer as respostas já corrigidas, ou seja, já codificadas como 0 (erro) e 1 (acerto), não havendo portanto necessidade de gabarito. No caso de esquemas amostrais complexos, pode-se fornecer ao programa pesos diferentes para cada um dos indivíduos. Assim como no PARSCALE, as informações devem estar em arquivos do tipo ASCII.

Caso seja fornecido o gabarito, é nessa fase (fase 1) que os itens são corrigidos, e são apresentadas também algumas estatísticas descritivas, como o número de sujeitos submetidos a cada item, número e porcentagem de acerto e algumas correlações de interesse, usadas na análise clássica. Essa etapa é de fundamental importância, pois além da oportunidade de conferir se os dados estão sendo lidos de maneira adequada, as estatísticas fornecidas servem como um diagnóstico preliminar dos itens com um ajuste ruim ou problemas no gabarito.

A fase 2 é de calibração dos itens, onde o usuário pode escolher os métodos de máxima verossimilhança marginal ou um método bayesiano de estimação por maximização da distribuição marginal a posteriori. O BILOG fornece ainda gráficos contendo informações de interesse, como as curvas características dos itens e de informação para cada item e para o teste global. Um atrativo do BILOG é que itens com correlações muito baixas são automaticamente retirados da análise, característica que não acontece com os demais programas. O software usa o valor de correlação bisserial inferior a $-0,15$ como padrão para retirada direta da calibração. Na fase 2 também é fornecido o teste de ajuste para cada item. São computados índices de ajuste aproximados por quiquadrado para cada item depois do ciclo final de estimação, onde a escala de habilidade é dividida em um número de intervalos sucessivos convenientes (máximo de 20) para mostrar as proporções da resposta. Logo após é computada uma razão de verossimilhança quiquadrado, de modo que a

frequência esperada não exceda 5. Os graus de liberdade são iguais ao número de intervalos combinados.

E a fase 3 é a fase de estimação das habilidades dos respondentes, baseadas nos resultados da fase anterior. Essas habilidades inicialmente são estimadas na escala de habilidade dos parâmetros. No entanto, pode-se especificar alguns tipos de mudanças na escala, que serão feitas tanto nas habilidades como nos parâmetros estimados na fase anterior. As habilidades podem ser estimadas pelo método de máxima verossimilhança, por esperança a posteriori (EAP) e por máximo a posteriori (MAP).

O BILOG também realiza a equalização dos itens, mas como esse assunto não foi abordado neste trabalho, não serão fornecidos detalhes sobre a mesma. As informações referentes à equalização encontram-se em Valle 1999.