

JULIANA OLIVEIRA GOMES



**ESTUDOS PSICOMÉTRICOS DA ESCALA BAPTISTA DE DEPRESSÃO – VERSÃO
ADULTO (EBADEP-A)**

ITATIBA

2013

JULIANA OLIVEIRA GOMES

ESTUDOS PSICOMÉTRICOS DA ESCALA BAPTISTA DE DEPRESSÃO – VERSÃO

ADULTO (EBADEP-A)

Conjunto de pesquisas apresentadas ao Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco e aprovadas em Arguição Final para fins de obtenção do Título de Doutor em Psicologia.

Apoio: CAPES

ITATIBA

2013

WM 190 Gomes, Juliana Oliveira.
G614e Estudos psicométricos da Escala Baptista de
depressão – versão adulto (EBADEP-A). / Juliana Oliveira
Gomes. -- Itatiba, 2013.
157 p.

Tese (doutorado) – Programa de Pós-Graduação
Stricto Sensu em Psicologia da Universidade São Francisco.
Orientação de: Makilim Nunes Baptista.

1. Depressão. 2. Comportamento autodestrutivo.
3. Avaliação psicológica. 4. Validade estatística. I. Baptista,
Makilim Nunes II. Título.

Ficha catalográfica elaborada pelas bibliotecárias do Setor de
Processamento Técnico da Universidade São Francisco.



UNIVERSIDADE SÃO FRANCISCO
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU*
EM PSICOLOGIA

Juliana Oliveira Gomes defendeu a tese "ESTUDOS PSICOMÉTRICOS DA ESCALA BAPTISTA DE DEPRESSÃO - VERSÃO ADULTO (EBADEP-A)" aprovada pelo Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco em 19 de fevereiro de 2013 pela Banca Examinadora constituída por:

Prof. Dr. Makilim Nunes Baptista
Presidente

Profa. Dra. Mirna Yamazato koda

Prof. Dr. Cláudio Garcia Capitão

Profa. Dra. Irani Iracema de Lima Argimon

Prof. Dr. Lélío Moura Lourenço

*O êxito da vida não se mede pelo caminho
que você conquistou, mas sim pelas
dificuldades que superou no caminho.*

Abraham Lincoln

Agradecimentos

Agradeço primeiramente a Deus, a quem sempre depus minha fé, mesmo nos momentos em que eu me sentia só, e sem coragem para continuar. E foram nos momentos mais difíceis que a fé me trouxe forças para dar os passos seguintes. Que eu possa também encontrar forças para os próximos. Amém.

Agradeço a minha família, sempre ao meu lado, mesmo quando a distância física é grande.

À minha mãe, Gracinha, por saber ouvir e ter sempre algo a dizer; à mãe Binha por ter sempre uma história feliz para contar. Minha irmã Dani e meu cunhado Aniel, pelo apoio e pelas conversas na cozinha. À minha sobrinha, Amanda, pelo som gostoso que invade meus ouvidos quando imita a irmã e me chama de “cocó”... À minha amada afilhada Maria Lúcia pelas brincadeiras, sorrisos e abraços apertados. Obrigada pelos desenhos e bilhetinhos, meu amor. Madrinha os guarda com muito carinho.

À minha família itatibense, Léia Vilerá, Eduardo Seixas, Victor Pereira, Anderson Lima, Camila Camilo, Robison Lima, Mariana Pinhatari e Carlos Collares. Obrigada pelo vinho, pela cachaça mineira, pelos filmes, risos, almoços, conversas, e por cada momento compartilhado.

Especialmente à Léia, a quem não só dividi o mesmo espaço de moradia, mas com quem construí um lar. Você me mostrou não só que é possível viver a intensidade do ser, mas também amar intensamente quem se é. Obrigada por estes olhos-cor-de-humor, pelo

sorriso bobo, parecido com soluço. Obrigada por cada abraço e cada bronca. Ganhei de Deus uma irmã de olhos claros!

À minha querida Marlene Alves, pela companhia ao longo dos anos, pela confiança em mim depositada. Nós nos encontramos no início de nosso mestrado e fomos alimentando uma amizade para a vida toda. Obrigada por acreditar em mim, até quando eu mesma não conseguia.

Ao meu anjo da guarda e afilhado Nelimar Ribeiro. Não sei usar palavras para lhe agradecer por tudo. Nada do que eu possa dizer é capaz de refletir o que sinto. Obrigada por me ensinar tanto e por gostar tanto dessa pessoa chata e sentimental que vos fala. Ao Roberto Luquini por me deixar entrar em sua vida e ainda levar mala! Obrigada pelo apoio, não só na tese, meu querido afilhado, mas em todos os aspectos da minha vida. Obrigada por compartilhar tanto amor e amizade.

Aos meus companheiros e amigos, ininterruptamente on-line, Fabio Hercules, Heitor de Almeida, Bruno Barbosa, Juliana Mendes e Alexandre Baldin, sempre prontos para ouvir meus lamentos e desabafos. Obrigada pela excelente companhia e pelos valiosos conselhos.

Aos meus amigos de longa data Mariana Corrêa, Luciana d'Ávila, Diêgo Candian e Alexandre Serpa. Vocês são mais que irmãos pra mim. Obrigada por acreditarem, por ajudarem sempre que eu desse o passo seguinte.

Agradeço aos meus professores, por estarem sempre disponíveis e por sempre ajudar, dar conselhos, palpites e sugestões.

Ao professor Makilim Baptista, pela orientação e amizade, desde o mestrado. Obrigada pela compreensão, pelo apoio e força. Você me ensinou coisas para além do acadêmico, tenha certeza disso.

Aos professores Fabián Rueda, Acácia Santos e Claudette Vendramini. Obrigada pela confiança de vocês. Obrigada por me ensinarem tanto.

Aos professores, integrantes da banca de Exame de Qualificação e Defesa, Prof^a Dr^a Mirna Koda, Prof Dr Cláudio Capitão, Prof^a Dr^a Irani Argimon e Prf Dr Lélío Lourenço, pelo carinho e disponibilidade. Agradeço pela leitura e as valiosas sugestões.

Aos meus companheiros de Laboratório, com quem convivi durante tanto tempo, Adriana Munhoz, Fernando Padovan, Hugo Cardoso, Marco Antonio e Gabriela Cremasco. Obrigada pela companhia, pela paciência, pelas leituras e sugestões. Obrigada também pela confiança em mim. Vocês são show!

Agradeço por fim aos estudantes que aceitaram responder aos questionários e participar das pesquisas, aos coordenadores de curso, os quais me receberam com muito carinho e atenção para a realização da pesquisa e à CAPES pelo apoio financeiro.

Resumo

Gomes, J. O. (2012). *Estudos Psicométricos da Escala Baptista De Depressão – Versão Adulto (EBADEP-A)*. Tese de Doutorado, Universidade São Francisco, Itatiba, SP.

Dentro do contexto principal de Avaliação Psicológica, o presente estudo procurou articular e discutir as qualidades psicométricas que um instrumento deve possuir para que seu uso seja adequado, a saber, a precisão, validade e também a normatização. Objetivou-se investigar as qualidades psicométricas da Escala Baptista de Depressão – Versão Adulto (EBADEP-A), tendo como base a Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI). Dentro dos objetivos específicos, procurou-se expor os procedimentos de adaptação e composição fatorial da Escala de Razões para Viver, versão brasileira da *Reasons for Living Scale* (RFL). A estrutura do presente trabalho foi dividida em quatro artigos. Nos três primeiros, focados na EBADEP-A, foram averiguadas evidências de validade baseada na relação com variáveis externas, do tipo convergente, e com base na estrutura interna, por meio de estudo fatorial confirmatório. Foram também realizados estudos de consistência interna com base no funcionamento diferencial do item, e normatização para o instrumento, pela análise de transferência de normas, pela TRI. O último artigo, destacando a RFL, foram apresentados os procedimentos de tradução da escala, e análise fatorial para a versão brasileira do instrumento. O RFL ficou composto por 28 itens, divididos em 3 fatores, “Crenças de sobrevivência e enfrentamento”, “Medo da morte e de desaprovação social” e “Motivação de controle externo”. Algumas limitações às pesquisas e sugestões para estudos futuros também são discutidos.

Palavras-chave: depressão; comportamento autodestrutivo; avaliação psicológica; validade estatística

Abstract

Gomes, J. O. (2012). *Psychometric Studies for the Baptista Depression Scale - Adult Version (EBADEP-A)*. Doctoral Thesis, São Francisco University, Itatiba, SP.

Within the context of Psychological Assessment, this study sought to articulate and discuss the psychometric qualities that an instrument should have so that their use is appropriate, namely the reliability, validity, and also the normalization. This study aimed to investigate the psychometric properties of the Baptista Depression Scale - Adult Version (EBADEP-A), based on Classical Test Theory (CTT) and Item Response Theory (IRT). Within the specific objectives one sought to expose the adaptation procedures and factorial composition of the Brazilian version of the Reasons for Living Scale (RFL). The structure of the present study was divided into four articles. In the first three, focused on EBADEP-A have been investigated evidences of validity based on the relationship with the external variables, converging type, and based on the internal structure through a factorial study confirmatory. Were also conducted internal consistency studies, based on differential item functioning, and standardization for the instrument, through a transfer of norms, via IRT. The last article, highlighting the RFL, were presented procedures for translation of the scale and factor analysis for the Brazilian version of the instrument. The RFL was composed of 28 items divided into three factors, "Beliefs of survival and coping", "Fear of death and of social disapproval" and "Motivation external control". Some limitations to the research and suggestions for future studies are also discussed.

Keywords: depression; self-destructive behavior; psychological assessment; statistical validity

Resumen

Gomes, J. O. (2012). Estudios psicométricos de la Escala de Depresión Baptista - Versión Adulto (EBADEP-A). Tesis de Doctorado, Postgrado en Psicología de la Universidad San Francisco, Itatiba, SP.

Dentro del contexto principal de la Evaluación Psicológica, este estudio trató de articular y discutir las cualidades psicométricas de un instrumento debe tener para que su uso sea apropiado, a saber, la fiabilidad, validez, y también la normalización. Este estudio tuvo como objetivo investigar las propiedades psicométricas de la Escala Baptista de Depresión - Versión Adulto (EBADEP-A), basado en la Teoría Clásica de los Tests (TCT) y la Teoría de Respuesta al Ítem (TRI). Entre los objetivos específicos buscó exponer los procedimientos de adaptación y la composición factorial de la Escala de Razones para Vivir, la versión brasileña de la *Reasons for Living Scale* (RFL). La estructura de este trabajo fue dividido en cuatro artículos. En los tres primeros, centrados en EBADEP-A se investigaron las evidencias de la validez basada en la relación con las variables externas, de tipo convergente, y en base a la estructura interna por medio de un estudio de confirmación factorial. Se realizaron estudios de consistencia interna basada en el funcionamiento diferencial del ítem, y la estandarización del instrumento, basados en transferencia de las normas, mediante el TRI. El último artículo, destacando el RFL fueron presentados los procedimientos de traducción de la escala y el análisis factorial de la versión brasileña del instrumento. El RFL estaba compuesto de 28 ítems divididos en tres factores: " Las creencias de supervivencia y coping", "Miedo a la muerte y la desaprobación social" y el "control de la motivación externa." Algunas de las limitaciones a la investigación y propuestas para futuros estudios también se discuten.

Palabras clave: depresión; conducta autodestructiva; evaluación psicológica; validez estadística

Sumário

| | |
|---|------|
| Lista de Figuras | xiii |
| Lista de Tabelas | xv |
| Lista de Anexos | xv |
| Apresentação | 1 |
| Artigo 1. Escala Baptista de Depressão – Versão Adulto: evidências de validade e consistência interna em amostra de estudantes..... | 17 |
| Introdução..... | 20 |
| Método..... | 25 |
| Participantes | 25 |
| Instrumentos | 25 |
| Procedimentos | 27 |
| Resultados..... | 28 |
| Discussão..... | 33 |
| Referências | 37 |
| Artigo 2. Análise Fatorial confirmatória da Escala Baptista de Depressão – Versão Adulto | 50 |
| Introdução..... | 53 |
| Método..... | 61 |
| Participantes | 61 |
| Instrumentos | 61 |
| Procedimentos | 62 |
| Resultados..... | 63 |
| Discussão | 66 |
| Referências | 72 |

Artigo 3. Normalization procedure for the Baptista Depression Scale - Adult Version

| | |
|--|----|
| (EBADEP-A): transferring of norms..... | 80 |
| Method..... | 88 |
| Participants | 88 |
| Instruments | 88 |
| Procedures | 89 |
| Results | 89 |
| Discussion..... | 93 |
| References | 97 |

Artigo 4. Tradução para o Português, Adaptação e Análise Fatorial da Reasons for Living

| | |
|--|-----|
| Scale | 109 |
| Introdução..... | 112 |
| Método..... | 120 |
| Participantes | 120 |
| Instrumentos | 120 |
| Procedimentos | 121 |
| Resultados..... | 122 |
| Estudo 1 | 123 |
| Estudo 2 | 124 |
| Estudo 3 | 126 |
| Discussão..... | 127 |
| Referências | 132 |
| Conclusões e Considerações Finais..... | 141 |
| ANEXOS | 152 |

Lista de Figuras

| | |
|---|-----|
| Artigo 1..... Figura 1. Gráfico de probabilidades de resposta em relação a cada opção de resposta da EBADEP-A | 49 |
| Artigo 2.Figura 1. Gráfico <i>Scree Plot</i> | 78 |
| Artigo 2. Figura 2. Modelo unifatorial da EBADEP-A D..... | 79 |
| Artigo 3. Figure 1. Distribution of response options of the instruments..... | 103 |

Lista de Tabelas

| | |
|---|-----|
| Artigo 1. Tabela 1. Sumário de informações sobre a EBADEP-A de acordo com o modelo de Rasch..... | 45 |
| Artigo 1. Tabela 2. Parâmetros de ajuste dos itens e das pessoas, com exclusão de casos extremos | 46 |
| Artigo 1. Tabela 3. Mapa de Itens e Pessoas | 47 |
| Artigo 1. Tabela 4. Itens com funcionamento diferencial por sexo, de acordo com a medida de contraste | 48 |
| Artigo 2. Tabela 1. Índices de ajuste dos dois modelos testados para a EBADEP-A | 77 |
| Artigo 3. Table 1. Map of Items for CES-D and EBADEP-A | 104 |
| Artigo 3. Table 2. Conversion of scores into Thetas for CES-D..... | 105 |
| Artigo 3. Table 3. Conversion of scores into Thetas for EBADEP-A..... | 106 |
| Artigo 3. Table 4. Crosstabulation between the diagnostic categories of the instruments | 108 |
| Artigo 4. Tabela 1. Matriz de componentes, comunalidades e alfas de Cronbach para cada fator e escala total | 140 |

Lista de Anexos

| | |
|--|-----|
| Anexo A – Termo de Consentimento Livre e Esclarecido | 153 |
| Anexo B – Parecer emitido pelo Comitê de Ética em Pesquisa | 154 |
| Anexo C – Questionário de Identificação..... | 155 |
| Anexo D – Escala de Razões para Viver (versão traduzida)..... | 156 |

Apresentação

Uma das áreas da Psicologia que pode ser relacionada a quaisquer campos de atuação da profissão é a Avaliação Psicológica (AP). O psicólogo que utiliza a AP em seu contexto clínico, independente de sua área de atuação, bem como a teoria que embasa seu trabalho, leva em conta que se trata de um processo amplo para entendimento de questões relativas ao funcionamento psíquico de uma pessoa, dentro de um determinado período de tempo e que pode auxiliá-lo na tomada de decisão (Noronha, 2009; Noronha & Alchieri, 2004).

São vários os autores que se propõem a sistematizá-la, e por esta razão, julga-se pertinente apresentar os principais conceitos disponíveis na literatura. Para Wechsler (1999) a AP tem o objetivo de obtenção de subsídios sobre as dimensões psicológicas de um indivíduo ou grupo, a fim de se conseguir um melhor entendimento do ser humano, norteando decisões e planejando intervenções. Já para Parpinelli e Lunardelli (2006), a avaliação é comandada pela coleta de dados e interpretação dos resultados obtidos das mesmas, por meio dos quais se podem estudar os métodos e instrumentos psicológicos.

De fato, trata-se de uma área muito criticada e discutida na Psicologia (Noronha, 2009). Durante muitos anos, a AP foi citada e pautada como sinônimo de aplicação de testes psicológicos, e por esta razão seu ensino foi durante muito tempo centrado no treinamento no manuseio dos instrumentos, tanto objetivos quanto projetivos (Noronha & Alchieri, 2004). Com o passar do tempo, esta problemática passou a ser discutida de forma mais extensa e a AP passou a ser entendida como um processo amplo e complexo, que vai além da simples aplicabilidade de suas técnicas específicas, mas que inicia antes mesmo da escolha do teste em si, e também inclui o estudo sistemático do que se pretende avaliar, as

interpretações e contextualizações dos achados clínicos, ajudando na tomada de decisão (Amarnani, 2009; Andriola, 1996; Noronha, 2009; Noronha & Alchieri, 2002; 2004).

Historicamente, o estudo da AP começou de forma restrita, limitada primeiramente a questões referentes às testagens de pessoas consideradas mentalmente “retardadas” ou com problemas mentais, e posteriormente ao desenvolvimento da inteligência humana (Anastasi & Urbina, 2000; Pasquali, 2001). Durante muito tempo diversos autores se centraram na tentativa de operacionalizar a conceituação de “Teste psicológico” (Anastasi & Urbina, 2000; Cronbach, 1996; Pasquali, 2001), o que talvez tenha contribuído para esta associação inicial minimalista de AP com aplicação de testes.

Ao lado disso, com a intensa divulgação dos testes de inteligência iniciada entre as décadas de 1910 e 1930, impulsionados provavelmente pelo impacto da I Guerra Mundial, vários foram os profissionais designados para desenvolver testes de seleção de recrutas. Esta exposição por um lado favoreceu a Psicologia, popularizando o uso de testes psicológicos. Mas, por outro lado, atrasou o avanço da Avaliação Psicológica, com o uso indiscriminado dos instrumentos construídos e disponibilizados (Ambiel & Pacanaro, 2011; Pasquali, 2001).

Por esta razão, o estudo de Psicologia necessitava de uma organização, e por influência do surgimento das Universidades entre 1932 e 1962, as questões envolvendo a Avaliação começaram a ser discutidas nos termos da tentativa de sistematização de seus métodos (Ambiel & Pacanaro, 2011; Pasquali, 2001). Isto se refletiu tanto na área de pesquisa quanto no exercício da profissão clínica, e por isso, pode-se chamar tal período como “fase de ouro” da construção dos instrumentos de medida (Andriola, 1996; Noronha & Alchieri, 2002).

A partir de 1951 pode ser observado um número crescente de artigos envolvendo avaliação da personalidade, desenvolvimento infantil, além de estudos de adaptação de instrumentos e instituição de normas e condutas éticas para a utilização dos testes psicológicos (Noronha & Alchieri, 2002; 2004). Assim, percebe-se o início de um movimento, especificamente no Brasil, para que o uso dos testes, e por consequência o estudo da AP, fosse retomado de forma consciente, contextualizado, e não de forma indiscriminada, como o contexto histórico mostrava até então.

Os anos entre 1962 e 1970 foram marcados pela regulamentação dos cursos de graduação em Psicologia e também da profissão de Psicólogo. Por um lado, tal marco foi benéfico na medida em que instituiu a AP como uma função privativa ao Psicólogo. Por outro lado, resultou novamente em um processo de lentidão no avanço dos estudos, pois com a popularização da Psicologia, e a grande procura por profissionais, a preocupação com a qualidade nas medidas utilizadas caiu, talvez pela ausência de pesquisas relacionadas, que pudessem nortear os usos dos instrumentos e interpretação dos resultados encontrados (Ambiel & Pacanaro, 2011; Andriola, 1996; Pasquali, 2001; Noronha, 2009; Noronha & Alchieri, 2002).

Como pôde ser visto, o estudo da AP, e conseqüentemente a organização do uso dos instrumentos de medida sofreu altos e baixos historicamente, e talvez tenha sido somente com a criação dos programas de Pós-Graduação em Psicologia, entre 1970 e 1990, que a preocupação com a qualificação não só dos testes mas também dos profissionais e o desenvolvimento de pesquisas tenha sido solidificada. A partir da década de 1990, identificada por Pasquali (2001) como o “Quinto período da Avaliação Psicológica no Brasil”, a problemática do aprimoramento do ensino em AP passou a ser notória.

A partir de então, não somente a divulgação como também o processo de construção dos instrumentos se tornou alvo de preocupação e interesse, no sentido de instigar o uso de testes confiáveis e a retomada da pesquisa nesta área (Andriola, 1996; Pasquali, 2001). Nesta nova perspectiva criada, classificada por Andriola (1996) como um momento de “maturidade da Psicologia”, podem ser observadas medidas concretas no sentido de sistematização e organização do uso de testes e também do estudo da AP.

Como exemplo desta nova fase da AP no Brasil, pode-se citar a divulgação da Resolução do Conselho Federal de Psicologia nº002/2003, na qual foram estabelecidos, com base em normas internacionais, os critérios para que um instrumento psicológico possa ser distribuído no Brasil e utilizado por Psicólogos (AERA, APA & NCME, 1999; CFP, 2003; ITC, 2000; Noronha, 2009). Tal resolução parte do princípio de que o uso correto de um instrumento adequado permite ao profissional investigar diferentes aspectos psicológicos, mesmos os não tão facilmente observáveis, como é o caso da Depressão e Ideação suicida, construtos priorizados na presente pesquisa (ITC, 2000; Noronha, 2009).

A Depressão era apontada pela Organização Mundial da Saúde como uma doença crônica e incapacitante (Brundtland, 2000; OMS, 2001). Ela se configura como construto complexo, com diferentes classificações, categorizações nosológicas e modelos explicativos, o que leva à necessidade da realização constante de estudos que possam acompanhar o seu surgimento, desenvolvimento, manutenção e avaliação (APA, 2002; Beck, Rush, Shaw, & Emery, 1997; Del Porto, 1999; Ferster, 1973; Holland & Skinner, 1969; Matos, Matos, & Matos, 2006; Skinner, 1974).

Considerando sua propriedade multifatorial, diferentes aspectos dentre fatores de risco e características associadas, como por exemplo, a idade, estado civil, e histórico

familiar de diagnóstico, são estudados, não somente relacionando-se às teorias, mas no que concerne ao observado em estudos epidemiológicos, de incidência, prevalência, e rastreamento (Bahls & Bahls, 2002; Baptista, 2004; Boyd, Diamond, & Bourjolly, 2006; Khandolwal, & cols., 2001; Peron, Neves, Brandão, & Vicentini, 2004; dentre outros). Dentre os diversos fatores, um dos que mais recebem destaque diz respeito às diferenças entre gêneros no transtorno depressivo, uma vez que as mulheres são apontadas como mais susceptíveis à doença que os homens (Angst, & cols., 2002; Aros & Yoshida, 2009; Baptista, 2004; Carrillo, Rojo, & Staats, 2004; Justo & Calil, 2006; Kessler, 2003; Nolen-Hoeksema, Larson, & Graysib, 1999; Sprock & Yoder, 1997).

Ao lado disso, a Ideação suicida, ou seja, o comportamento suicida caracterizado por pensamentos recorrentes de morte, acompanhado do desejo pelo suicídio, com ou sem uma tentativa em período recente (Beck, Kovacs & Weissman, 1979; Beck, Steer & Ranieri, 1988) é de alta incidência mundial, o que levou a Organização Mundial de Saúde a classificá-lo como um problema de saúde pública (OMS, 2001; 2006; WHO, 1999). Por outro lado, este construto pode ser elemento fundamental no que diz respeito a implementação de ações preventivas em diversas populações (Werlang, Borges & Fensterseifer, 2005), entretanto, por outro lado, enquanto os estudos sobre o comportamento suicida apresentam grande número de publicações, principalmente as de caráter epidemiológico, o mesmo não ocorre para a ideação suicida, a qual recebe menor destaque, talvez pela sua característica mais subjetiva, mais relacionada a pensamentos e crenças que a comportamentos observáveis (Nock, & cols., 2008; Werlang, Borges & Fensterseifer, 2005).

Dentro do contexto específico da Avaliação Psicológica da Depressão e Ideação Suicida, principalmente no Brasil, embora haja um vasto leque de opções para avaliação do primeiro, seja por meio de entrevistas diagnósticas, escalas de avaliação do pesquisador ou de autoavaliação (Calil & Pires, 1998; Moreno & Moreno, 2005), para o segundo, somente um instrumento possui as qualidades psicométricas requisitadas para que ele possa ser utilizado no contexto clínico brasileiro (CFP, 2011; Cunha & Werlang, 1996; Cunha, 2001), o que de certa forma limita a quantidade de publicações e/ ou informações divulgadas.

Neste momento, dois pontos importantes devem ser destacados, inicialmente, em relação à depressão, e posteriormente, ao suicídio. Embora possam ser encontrados diversos meios de mensuração, a análise dos sintomas priorizados pelas escalas de depressão disponíveis, com as devidas qualidades psicométricas investigadas, apresenta uma tendência a considerar um ou outro grupo de sintomas, em detrimento aos demais (Calil & Pires, 1998; Moreno & Moreno, 2005). Além disso, considerando o suicídio como um construto delicado de se avaliar, por envolver questionamentos e pensamentos em que as pessoas podem não se sentir confortáveis em declarar (Kaplan & cols., 1994), são apresentadas duas escalas, uma para avaliação da depressão e a outra, de ideação suicida, as quais possuem características que respondem bem a estas indagações.

A Escala Baptista de Depressão – versão Adulto (EBADEP-A), construída no Brasil, e aprovada pelo Sistema de Avaliação de Testes Psicológicos (CFP, 2011) foi organizada de modo a distribuir, forma equilibrada, as categorias sintomáticas as quais avalia, quais sejam, sintomas de humor, vegetativos ou somáticos, motores, sociais, cognitivos, de ansiedade e irritabilidade (Baptista, 2012; Calil & Pires, 1998). Calil e Pires

(1998) apresentaram a distribuição destas categorias de sintomas de escalas mundialmente conhecidas, como é o caso do Inventário de Depressão de Beck (BDI), por exemplo, o qual se constitui com 52% de itens cognitivos, seguidos de 29% vegetativos, o que já praticamente completa o escopo possível do instrumento.

Outros exemplos que podem ser citados são a escala de autoavaliação de Zung, a qual possui 35% de seus itens de caráter vegetativo e 35% cognitivos, e a *Center for Epidemiologic Studies - Depression* (CES-D), a qual, desde sua construção, foi descrita contendo foco em questões afetivas e humor depressivo (Calil & Pires, 1998; Radloff, 1977). A Escala Baptista de Depressão – versão Adulto, por sua vez, foi apresentada por Baptista (2012) contendo 33% de descritores sintomáticos do tipo cognitivo, 20% de humor, 18% de sintomas sociais, 18% vegetativos ou somáticos, 4,5% motores, 4,5% de irritabilidade e 2% de ansiedade.

No que diz respeito à ideação suicida, o instrumento apresentado no presente estudo a *Reasons for Living Scale* (RFL), escala a qual se propõe a mensurar, de maneira indireta, ideação suicida, por meio de questões relativas a diferentes motivos para que a pessoa não apresente o comportamento suicida (Linehan, & cols., 1983; Osman, & cols., 1993; 1999). Este teste se utilizou da Teoria Cognitiva da Depressão (Beck, Rush, Shaw, & Emery, 1997) como base norteadora para a construção de seus itens, e por esta razão considera a importância do sistema de crenças pessoais no estudo do suicídio. Assim, sua hipótese principal centra-se na concepção de que os indivíduos vulneráveis ao comportamento suicida apresentam poucas crenças relativas a motivos para estar vivo (Beck, Steer, & Ranieri, 1988; Linehan, & cols., 1983; Osman, & cols., 1993; 1999).

Por um lado, a *Reasons for Living Scale* (RFL) pode se configurar como uma boa alternativa ao estudo da ideação suicida, mencionada anteriormente como um construto de delicada mensuração (Kaplan, & cols., 1994). Por ser um instrumento de autoavaliação, sua utilização pode ser vantajosa, quando comparada a entrevistas clínicas, não somente por ser de rápida duração e pelo caráter *screening* de sintomas, mas por não ser necessária a participação ativa do profissional aplicador durante as respostas às questões, o que pode auxiliar com que a pessoa não se sinta desconfortável, como poderia ocorrer em uma entrevista verbal. Por outro lado, trata-se de um instrumento internacional, o qual não deve ser livremente traduzido para sua utilização no contexto brasileiro. São necessários diferentes estudos ao se introduzir à população um instrumento oriundo de outro país, afinal, não somente as diferenças de linguagem estão envolvidas, mas também culturais e contextuais (Manzi-Oliveira, Balarini, Marques, & Pasian, 2011).

Isto posto, o objetivo principal do presente estudo versou na investigação das qualidades psicométricas básicas da Escala Baptista de Depressão – versão adulto (EBADEP-A), considerando, dentro dos objetivos específicos, a tradução e validação da *Reasons for Living Scale* (RFL) para o português do Brasil. Para que tais objetivos pudessem ser alcançados teórica e metodologicamente, a estrutura do trabalho foi dividida em quatro artigos, os três primeiros centrados na depressão, especificamente na EBADEP-A, e por fim, um artigo focado na Escala de Razões para Viver, versão brasileira da RFL.

Ressalta-se, porém, que embora os artigos sejam apresentados de forma independente, os procedimentos metodológicos dos estudos envolvidos foram conduzidos com base em um único projeto de pesquisa. Dessa forma, após a aprovação do Comitê de Ética em Pesquisa da Universidade São Francisco (Protocolo CAE 0422.0.142.000-11), os

instrumentos envolvidos, quais sejam, um questionário de identificação, a Escala Baptista de Depressão – Versão Adulto (EBADEP-A), a *Center for Epidemiologic Studies - Depression* (CES-D) e a *Reasons for Living Scale* (RFL), os quais foram devidamente descritos em cada um dos trabalhos empíricos, foram aplicados de forma conjunta, em grupos de 20 a 40 participantes por vez.

No primeiro artigo, sobre depressão, foram expostas as informações básicas para o estudo do transtorno, como conceituação, e classificação, de acordo com os manuais diagnósticos. Posteriormente, a escala de rastreamento epidemiológico *Center for Epidemiologic Studies* (CES-D) foi indicada como base para a investigação de evidências de validade baseada na relação com construtos relacionados para a EBADEP-A. Foram realizadas correlações com e sem controle de sexo entre os escores dos instrumentos, além de análises de diferenças entre as pontuações médias de homens e mulheres às escalas. Com base no modelo de Rasch da Teoria da Resposta ao Item, foram encontrados índices muito satisfatórios de consistência interna para a EBADEP-A, por meio do estudo do funcionamento diferencial do item.

Para o segundo artigo, foi exibido de maneira mais completa como a depressão tem sido estudada e avaliada, tanto no que se refere aos manuais diagnósticos internacionais, quanto pelos instrumentos reconhecidos pela comunidade científica. Especificamente sobre a Escala Baptista de Depressão, foram apresentadas duas possíveis estruturas fatoriais, uma baseada nos estudos iniciais do instrumento, a partir do qual se estabeleceram uma estrutura bifatorial, e a outra baseada no instrumento finalizado e publicado, de organização unifatorial. Foi realizada Análise Fatorial Confirmatória para a EBADEP-A, sendo apreciadas ambas as organizações supracitadas. Os índices de ajuste foram verificados e

comparados, e a interpretabilidade dos índices, juntamente com critérios semânticos, sugeriu a estrutura unifatorial como mais adequada.

O terceiro artigo, por sua vez, com intencionalidade de publicação no exterior, foi estruturado em língua inglesa. Sua organização deu-se de forma a apresentar teoricamente as principais qualidades psicométricas de um instrumento psicológico, validade e precisão, bem como a normatização, foco do estudo. Dentro da perspectiva da Teoria da Resposta ao Item, foi estruturado um meio de normatização da EBADEP-A, com base na pontuação de corte da *Center for Epidemiologic Studies – Depression Scale* (CES-D). Assim, as medidas de ambos os instrumentos foram calibradas de forma a permitir a comparação e transferência do padrão normativo de um instrumento ao outro, por meio da equalização.

Por fim, a ideação suicida e as razões para viver foram contempladas no último artigo, juntamente com os conceitos básicos norteadores para o estudo do comportamento suicida. O artigo argumenta, de forma breve, a importância do estudo da ideação suicida de forma indireta, devido à escassez de instrumentos brasileiros que realizem tal tarefa, e apresenta os procedimentos de tradução, adaptação inicial e análise fatorial da Escala de Razões para Viver (do Inglês, *Reasons for Living Scale*). O objetivo do trabalho foi investigar evidências de validade baseadas no conteúdo e baseadas na estrutura interna da versão traduzida da escala, bem como identificar a organização da mesma frente às amostras brasileiras. A estrutura fatorial mais ajustada estatística e semanticamente foi a de três fatores, nomeados “Crenças de sobrevivência e enfrentamento”, “Medo da morte e de desaprovação social” e “Motivação de controle externo”, sendo encontrados índices satisfatórios para as evidências de validade iniciais do instrumento.

Referências

- Amarnani, R. (2009). Two Theories, One Theta: A Gentle Introduction to Item Response Theory as an Alternative to Classical Test Theory. *The International Journal of Educational and Psychological Assessment*, 3, 104-109.
- Ambiel, R. A. M., & Pacanaro, S. V. (2011). Da testagem à Avaliação Psicológica: aspectos históricos e perspectivas futuras. Em R. A. M. Ambiel, I. S. Rabelo, S. V. Pacanaro, G. A. S. Alves, I. F. A. S. Leme (Orgs). *Avaliação Psicológica: guia de consulta para estudantes e profissionais de psicologia*. São Paulo: Casa do Psicólogo.
- American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME]. (1999). *Standards for Psychology and Educational Testing*. Washington, DC: American Psychology Association.
- American Psychiatric Association [APA]. (2002). *DSM-IV-TR - Manual diagnóstico e estatístico de transtornos mentais*. (4ª ed). Porto Alegre: Artmed.
- Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica*. Veronese, M.A.V. (trad.). Porto Alegre: Artmed Editora.
- Andriola, W. B. (1996). Avaliação Psicológica no Brasil: considerações a respeito da formação dos psicólogos e dos instrumentos utilizados. *Psique*, 6(2), 99-108.
- Angst, J., Gamma, A., Gastpar, M., Lépine, J-P., Mendlewicz, J., & Tylee, A. (2002). Gender differences in depression. *European Archives of Psychiatry and Clinical Neuroscience*, 252, 201-209.
- Aros, M. S., & Yoshida, E. M. P. (2009). Estudos da depressão: instrumentos de avaliação e gênero. *Boletim de Psicologia*, 59(130), 61-76.

- Bahls, S-C., & Bahls, F. R. C. (2002). Depressão na adolescência: características clínicas. *Interação em Psicologia*, 6(1), 49-57.
- Baptista, M. N. (2012). *Manual Técnico da Escala Baptista de Depressão – Versão Adulto (EBADEP-A)*. São Paulo: Vetor.
- Baptista, M. N. (2004). *Suicídio e Depressão: Atualizações*. São Paulo: Guanabara Koogan.
- Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of Suicidal Intention: The Scale for Suicide Ideation. *Journal of Consulting and Clinical Psychology*, 47(2), 343-352.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1997). *Terapia da Depressão*. Rio de Janeiro: Zahar.
- Beck, A. T., Steer, R. A., & Ranieri, W. F. (1988). Scale for suicide ideation: psychometric properties of a self-report version. *Journal of Clinical Psychology*, 44(4), 499-505.
- Boyd, R. C., Diamond, G. S., & Bourjolly, J. N. (2006) Developing a Family-Based Depression Prevention Program in Urban Community Mental Health Clinics: A Qualitative Investigation. *Family Process*, 45(2), 187-203.
- Brundtland, G. H. (2000). The Mental Health in XXI Century. *Bulletin of the World Health Organization*, 78(4), 411.
- Calil, H. M., & Pires, M. L. N. (1998). Aspectos gerais das escalas de avaliação de depressão. *Revista de Psiquiatria Clínica*, 25(5), 240-244.
- Carrillo, J. M., Rojo, N., & Staats, A. W. (2004). Women and vulnerability to depression: some personality and clinical factors. *The Spanish Journal of Psychology*, 7(1), 29-39.
- Conselho Federal de Psicologia [CFP]. (2011). Sistema de Avaliação de Testes Psicológicos (SATEPSI). Disponível em: <http://www.pol.org.br/satepsi/>

- Conselho Federal de Psicologia – CFP (2003). Resolução nº 002/2003 [On-line]. Recuperado de http://www.pol.org.br/pol/export/sites/default/pol/legislacao/legislacaoDocumentos/resolucao2003_02.pdf
- Cronbach L. J. (1996). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artes Médicas.
- Cunha, J. A. (2001). *Manual da versão em português das Escalas de Beck*. Casa do Psicólogo, São Paulo.
- Cunha, J. A., & Werlang, B. C. (1996). Um estudo com a Escala de Desesperança de Beck em grupos clínicos e não-clínicos. *Psico* 27, 189-197.
- Del Porto, J. A. (1999). Conceito e Diagnóstico. *Revista Brasileira de Psiquiatria*, 21 (1), 6-11.
- Ferster, C. B. (1973). A functional analysis of depression. *American Psychologist*, 28, 857-870
- Holland, J. G., & Skinner, B. F. (1969). *A Análise do Comportamento*. São Paulo: E.P.U.
- International Test Commission – ITC (2000). Guidelines on Adapting Test. *International Test Commission*. [On-Line]. Recuperado de <http://www.intestcom.org/upload/sitefiles/40.pdf>
- Justo, L. P., & Calil, H. M. (2006). Depressão – o mesmo acometimento para homens e mulheres? *Revista de Psiquiatria Clínica*, 33(2), 74-79.
- Kaplan, M. L., Asnis, G. M., Sanderson, W. C., Keswani, L., Lecuosa, J. M., & Joseph, S. (1994). Suicide assessment: clinical interview vs. self-report. *Journal of Clinical Psychology*, 50(2), 294-298.

- Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders, 74*, 5-13.
- Khandolwal, S., Chowdhury, A., Regmi, S. K., Mendis, N., & Kittirattanapaiboon, P. (2001). Conquering depression: you can get out of the blues. *World Health Organization, Regional Office for South-East Asia*, Nova Delhi.
- Linehan, M. M., Goodstein, J. L., Nielsen, S. L., & Chiles, J. A. (1983). Reasons for Staying Alive. When you are thinking of killing yourself. The Reasons for Living Inventory. *Journal of Consulting and Clinical Psychology, 51*(2), 276-286.
- Manzi-Oliveira, A. B., Balarini, F. B., Marques, L. A. S., & Paisan, S. R. (2011). Adaptação cultural de instrumentos de Avaliação Psicológica: levantamento de estudos realizados no Brasil de 2000 a 2010. *Psico-USF, 16*(3), 367-381.
- Matos, E. G., Matos, T. M. G., & Matos, G. M. G. (2006). Depressão Melancólica e Depressão Atípica: aspectos clínicos e psicodinâmicos. *Estudos de Psicologia, 23*(2), 173-179.
- Moreno, R. A., & Moreno, D. H. (1998). Escalas de depressão de Montgomery & Åsberg (MADRS) e de Hamilton (HAM-D). *Revista de Psiquiatria Clínica, 25*(5), 1-17.
- Nock, M. K., Borges, G., Bromet, E. J., Cha, C. B., Kessler, R. C., & Lee, S. (2008). Suicide and Suicidal Behavior. *Epidemiologic Reviews, 30*, 133-154.
- Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the Gender Difference in Depressive Symptoms. *Journal of Personality and Social Psychology, 77*(5), 1061-1072.

- Noronha, A. P. P. (2009). Testes Psicológicos: conceito, uso e formação do psicólogo. Em Claudio S. Hutz (org). *Avanços e polêmicas em avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Noronha, A. P. P., & Alchieri, J. C. (2004). Conhecimento em Avaliação Psicológica. *Estudos de Psicologia*, 21(1), 43-52.
- Noronha, A. P. P., & Alchieri, J. C. (2002). Reflexões sobre os Instrumentos de Avaliação Psicológica. Em Ricardo Primi (org.). *Temas em avaliação psicológica*. Campinas: Imprensa Digital do Brasil
- Organização Mundial da Saúde – OMS (2001). *Relatório Mundial de Saúde*. Saúde Mental: nova concepção, nova esperança. [On-Line]. Recuperado de http://www.who.int/whr/2001/en/whr01_djmessage_po.pdf
- Osman, A., Giffords, J., Jones, T., Lickiss, L., Osman, J., & Wenzel, R. (1993). Psychometric Evaluation of the Reasons for Living Inventory. *Psychological Assessment*, 5(2), 154-158.
- Osman, A., Kopper, B. A., Linehan, M. M., Barrios, F. X., Gutierrez, P. M., & Bagge, C. L. (1999). Validation of the Adult Suicidal Ideation Questionnaire and the Reasons for Living Inventory in an Adult Psychiatric Inpatient Sample. *Psychological Assessment*, 11(2), 115-123.
- Parpinelli, R. F., & Lunardelli, M. C. F. (2006). Avaliação psicológica em processos seletivos: contribuições da abordagem sistêmica. *Estudos de Psicologia*, 23(4), 463-471.
- Pasquali, L. (2001). *Técnicas de Exame Psicológico – TEP: fundamentos das técnicas psicológicas*. São Paulo: Casa do Psicólogo.

- Peron, A. P., Neves, G. Y. S., Brandão, M., & Vicentini, V. E. P. (2004) Aspectos biológicos e sociais da depressão. *Arquivos de Ciência da Saúde da Unipar, Umuarama*, 8(1), 45-48.
- Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, 1(3), 385-401.
- Skinner, B. F. (1974). *Ciência e Comportamento Humano*. São Paulo: EDART.
- Sprock, J., & Yoder, C. Y. (1997). Women and Depression: an update on the report of the APA Task Force. *Sex Roles*, 36(516), 269-303.
- Urbina, S. (2004). *Essentials of Psychological Testing*. New Jersey: John Wiley & Sons, Inc.
- Wechsler, S. M. (1999). Guia de procedimentos éticos para a avaliação psicológica. Em S. M. Wechsler & R. S. L. Guzzo (Orgs). *Avaliação psicológica: perspectiva internacional* (pp. 133-141). São Paulo: Casa do Psicólogo.
- Werlang, B. S. G., Borges, V. R., & Fensterseifer, L. (2005). Fatores de risco ou proteção para a presença de ideação suicida na adolescência. *Revista Interamericana de Psicología*, 39(2), 259-266.
- World Health Organization (1999). *Figures and Facts about Suicide*. Geneva: WHO.

Artigo 1

Escala Baptista de Depressão – Versão Adulto: evidências de validade e consistência interna em amostra de estudantes

O presente artigo teve como objetivo a análise de evidências de validade e consistência interna para a Escala Baptista de Depressão – Versão Adulto (EBADEP-A). Participaram 581 estudantes com média de idade de 24,20 anos ($DP = 8,76$), majoritariamente mulheres (74,5%), os quais responderam à EBADEP-A, à *Center for Epidemiologic Studies - Depression* (CES-D). As correlações encontradas, bem como as diferenças de pontuações médias indicaram importantes evidências de validade baseada na relação com variáveis externas para a EBADEP-A. Quanto à consistência interna, observaram-se bons índices de ajuste de itens e pessoas ao modelo de Rasch, bem como áreas delimitadas entre as opções de resposta da escala, não tendo sido encontrados índices com alta incidência de ruídos. Quanto ao funcionamento diferencial do item em relação a sexo, quatro itens apresentaram viés de resposta ao grupo feminino, enquanto três foram beneficiados pelo grupo masculino, entretanto, pelo princípio de equidade, pode-se concluir um equilíbrio nos vieses, e consistência interna satisfatória para a EBADEP-A.

Palavras-chave: validade estatística; precisão do teste; teoria de resposta ao item

Baptista Depression Scale - Adult Version: evidence of validity and internal consistency in university samples.

This article aims to analyze the validity and internal consistency for the Baptista Depression Scale - Adult Version (EBADEP-A). Participants were 581 students with a mean age of 24.20 years ($SD = 8.76$), mostly women (74.5%), which answered to the EBADEP-A and the Center for Epidemiologic Studies - Depression (CES-D). The correlations and the differences in mean scores indicated significant evidence of validity based on the relationship with external variables to the EBADEP-A. Regarding the internal consistency, there were observed good fit indices of persons and items to the Rasch model, as well as areas bounded between the response options of the scale, not having been found high incidence rates of noise. As for differential item functioning in relation to sex, four items showed response bias to the female group, while three were benefited by the male group, however, the principle of equity it can be concluded a biases balance, and satisfactory internal consistency for EBADEP-A.

Keywords: statistical validity; test precision; item response theory

Escala Baptista de Depresión – Versión de Adultos: evidencias de validez y consistencia interna en muestras de universitarias

Este artículo tiene como objetivo analizar la validez y consistencia interna de la Escala de Depresión Bautista - Versión Adulto (EBADEP-A). Los participantes fueron 581 estudiantes con una edad media de 24,20 años (DE = 8,76), en su mayoría mujeres (74,5%) que respondieron a EBADEP-A y la *Center for Epidemiologic Studies - Depression* (CES-D). Las correlaciones y las diferencias en las puntuaciones medias indican una evidencia significativa de validez basada en la relación con las variables externas a EBADEP-A. En cuanto a la consistencia interna, observaron buenos índices de ajuste de las personas y objetos en el modelo de Rasch, así como áreas delimitadas entre las opciones de respuesta de la escala, no habiéndose encontrado con altas tasas de incidencia de ruido. Con respecto al funcionamiento diferencial de los ítems en relación con el sexo, cuatro ítems mostraron parcialidad en la respuesta al grupo de mujeres, mientras que tres fueron beneficiados por el grupo masculino, sin embargo, por el principio de equidad, se puede concluir que existe un equilibrio en el sesgo y la fiabilidad satisfactoria a EBADEP-A

Palabras clave: validez estadística; prueba de precisión; la teoría de respuesta al ítem

Introdução

Diversas são as pesquisas citando a depressão como sendo uma doença bastante frequente mundialmente (Aros & Yoshida, 2009; Coutinho, Gontiès, Araújo, & Sá, 2003; Fleck, & cols., 2003; Matos, Matos, & Matos, 2006; OMS, 2001; Suija, Kalda, & Maaros, 2009; Vicente, Rioseco, Saldivia, Kohn, & Torres, 2005; Vismari, 2004). Em um importante relatório oficial sobre a saúde no mundo, a Organização Mundial de Saúde (OMS, 2001), apresentou-a como sendo a quarta maior causa de incapacitação, sendo esperado para 2020 que ela passe a ocupar o segundo lugar em países desenvolvidos, e o primeiro em países em desenvolvimento.

De maneira geral, diferentes maneiras de se explicar o transtorno depressivo são encontradas, e, por conseguinte, as formas como ele é operacionalizado e estruturado. Por esta razão, é importante considerar o aspecto multifatorial do transtorno, não limitando seu estudo à análise de suas possíveis causas, e dos processos e mecanismos envolvidos em seu desencadeamento, como é explicado pelos modelos e teorias. Faz-se imperativo ponderar outros motivos os quais podem estar envolvidos tanto no surgimento quanto manutenção da doença, principalmente possíveis influências ambientais que possam circundar uma pessoa ao longo de seu desenvolvimento. Neste sentido, preza-se sobre o estudo dos aspectos associados ao transtorno e fatores de risco (Baptista, 2004).

Um dos aspectos mais estudados no que diz respeito aos fatores de risco tem sido as diferenças entre gêneros (Angst, & cols., 2002; Aros & Yoshida, 2009; Baptista, 2004; Carrillo, Rojo, & Staats, 2004; Justo & Calil, 2006; Kessler, 2003; Nolen-Hoeksema, Larson, & Graysib, 1999; Sprock & Yoder, 1997). De acordo com o *National Institute of Mental Health* (NIMH, 2000), quando se estabelece o número de pessoas acometidas por

depressão por ano, pode-se observar que quase o dobro é do sexo feminino. Ainda, aproximadamente 20% das mulheres apresentaram algum episódio depressivo em algum momento da vida.

As discussões das possíveis causas para esta diferença acentuada giram em torno tanto de questões biológicas e hormonais, naturais ao desenvolvimento feminino, quanto a fatores ambientais e culturais (Angst, & cols., 2002; Carrillo, Rojo, & Staats, 2004; Kessler, 2003; NIMH, 2000; Piccinelli & Wilkinson, 2000). Tais diferenças podem refletir na forma como os registros são estabelecidos, e conseqüentemente na interpretação dos dados (Baptista, 2004).

No que diz respeito a avaliação da depressão, vários instrumentos podem ser encontrados, relativos a diferentes grupos amostrais e idades (Calil & Pires, 1998; Coutinho, Carolino & Medeiros, 2008; Gorestein & Andrade, 1998; Marcolino & cols, 2007; Moreno & Moreno, 2005; Santor, Gregus & Welch, 2006; Santos & cols, 2007). De acordo com Santor, Gregus e Welch (2006) o Inventário de Depressão de Beck (BDI), e *Center for Epidemiologic Studies* (CES-D) são as escalas de autoaplicação mais utilizadas internacionalmente.

No Brasil, Calil e Pires (1998) apresentaram como as escalas de autoavaliação mais utilizadas o BDI e a Escala de Depressão de Zung, com base em uma pesquisa sobre os aspectos gerais das escalas de depressão. Por sua vez, Baptista, Argimon e Yoshida (2012) em um estudo de revisão sistemática sobre instrumentos de avaliação psicológica em pesquisas no Brasil, indicaram o BDI o teste mais utilizado, independente do público alvo dos instrumentos. Quando consideradas as escalas de autoaplicação, direcionados a adultos, o BDI e o CES-D são as mais citadas.

A *Center for Epidemiologic Studies* (CES-D) é uma escala com objetivo de rastreamento de sintomatologia depressiva, com ênfase nos aspectos afetivos, e no humor deprimido (Radloff, 1977). Ela foi construída tendo como base outras escalas, a saber, Escala de Depressão de Zung, Inventário de Depressão de Beck (BDI), Escala Auto-aplicável de Raskin, Escala de Depressão do *Minnesota Multiphasic Personality Inventory* (MMPI), e Escala de Depressão de Gardner, podendo ser utilizada tanto no contexto clínico quanto não-clínico. O instrumento traduzido para o contexto brasileiro por Silveira e Jorge (1998) e possui diferentes evidências de validade, envolvendo contextos distintos (Batistoni, Neri, & Cupertino, 2007; 2010; Hauck Filho & Teixeira, 2011).

Entretanto, percebe-se que a CES-D apresenta uma ênfase nos sintomas obrigatórios da depressão, especialmente os componentes afetivos e humor depressivo (Radloff, 1977), o que pode levar a um levantamento incompleto de sintomas no rastreamento. Por esta razão, destaca-se a Escala Baptista de Depressão – Versão Adulto (EBADEP-A), instrumento brasileiro o qual apresenta uma proposta de distribuição mais ampla de descritores depressivos para a avaliação do construto (Baptista, 2012).

A EBADEP-A (Baptista, 2012) foi construída no Brasil, com base dos manuais diagnósticos internacionais DSM-IV-TR (APA, 2002) e CID-10 (OMS, 1993), na Teoria Cognitiva de Beck (Beck, Rush, Shaw, & Emery, 1997) e na Teoria Comportamental de Ferster (Ferster, Culbertson, & Boren, 1977) e pode ser utilizada em amostras clínicas e não-clínicas. Ao longo de sua construção, diferentes versões podem ser encontradas, também com o nome de Escala de Depressão (EDEP), e diferentes estudos de evidências de validade (Baptista, 2012; Baptista, Cardoso, & Gomes, 2012; Baptista & Carneiro, 2011;

Baptista & Gomes, 2011; Baptista, Souza, & Alves, 2008; Gomes & Baptista, 2010; dentre outros).

Pode-se perceber que ambas as escalas possuem diferentes pesquisas relativas às suas qualidades psicométricas básicas, as quais envolvem a validade e a precisão (Noronha & Alchieri, 2004; Noronha & Vendramini, 2003; Pasquali, 1999; Urbina, 2004). A validade diz respeito ao conjunto de evidências que tornam possíveis que os escores dos testes sejam determinantes das características que eles se propõem a medir (Cronbach, 1996; Urbina, 2004). Trata-se de um processo dinâmico e contínuo no qual instrumento deve apresentar resultados respaldados ao encontrado em outras pesquisas do mesmo tema, ou temas semelhantes (Primi, Muniz, & Nunes, 2009; Urbina, 2004).

Não se pode determinar um tipo específico e único de validade para um instrumento, mas sim investigar diferentes evidências, conforme contextos determinados, uma vez que a validação não decai sobre o teste em si, mas às interpretações teoricamente embasadas dos resultados gerados (Urbina, 2004). Assim, diferentes classificações possíveis para um instrumento são apresentadas, a saber, relacionadas ao conteúdo, baseadas no processo de resposta, na consequência de testagem, baseadas na estrutura interna, e na relação com outras variáveis, ou variáveis externas (AERA, APA, & NCME, 1999; Urbina, 2004).

Dentro do ponto de vista da Teoria Clássica dos Testes (TCT), fundamentado nas características da amostra, as evidências de validade baseada na relação com variáveis externas são analisadas a partir dos padrões de correlação encontrados entre os escores dos testes. Assim, ao se comparar instrumentos os quais avaliam os mesmos construtos, ou

relacionados, são esperadas correlações de alta magnitude (AERA, APA & NCME, 1999; Primi, Muniz, & Nunes, 2009).

Por sua vez, a precisão, confiabilidade, ou ainda fidedignidade, refere-se à estabilidade de medida do teste, ao quanto a análise realizada pelo instrumento se aproxima do real possível às características mensuradas (Pasquali, 1998; 1999; Urbina, 2004). Dentro desta perspectiva, o teste deve possuir consistência nas pontuações geradas e uma das formas de se atingir este objetivo se dá por meio da Teoria do Traço Latente ou Teoria da Resposta ao Item (TRI), com base no estudo do funcionamento diferencial do item (DIF) (Hambleton & Jones, 1993; Nunes & Primi, 2009; Pasquali & Primi, 2003; Valentini & Laros, 2011).

A TRI é formada por um conjunto de modelos com a finalidade de representação dos parâmetros do teste, focando nos itens e em suas características e na probabilidade de uma pessoa em dar uma determinada resposta, a partir do seu nível de habilidade, denominado *Theta* (Amarnani, 2009; Embretson & Reise, 2000; Pasquali & Primi, 2003; Rueda, 2007; Valentini & Laros, 2011). Na análise do DIF, os parâmetros dos itens podem ser estimados de forma independente da amostra avaliada, sendo possível verificar a existência de viés de resposta ao instrumento, e conseqüentemente a sua consistência interna (Nunes & Primi, 2009; Valentini & Laros, 2011).

Dessa forma, o presente estudo teve como propósito geral investigar as qualidades psicométricas da Escala Baptista de Depressão – Versão Adulto (EBADEP-A) em amostras de estudantes, por meio da Teoria Clássica dos Testes (TCT) e Teoria da Resposta ao Item (TRI). De maneira específica, objetivou-se, por meio da TCT, buscar evidências de validade baseada na relação com variáveis externas para a EBADEP-A, a partir da *Center*

for Epidemiological Studies – Depression (CES-D) e realizar análises de consistência interna e funcionamento diferencial do item em relação a sexo, por meio da TRI.

Método

Participantes

A amostra, escolhida por conveniência, foi composta por 581 estudantes de diferentes cidades interioranas de São Paulo (Bauru, Itatiba e Lins) e Minas Gerais (Caratinga, Juiz de Fora e Viçosa), sendo 148 (25,5%) homens e 433 (74,5%) mulheres. Do total de participantes, 29 (5%) eram alunos de cursos profissionalizantes, enquanto o restante se dividiu entre diversos cursos de graduação, como por exemplo, Enfermagem, Geografia, Psicologia, Direito, Biologia, além de dois cursos de pós-graduação, em Saúde Mental e Psicologia do Trânsito. Quanto ao estado civil, 471 (81,1%) eram solteiros, variando entre 18 e 63 anos ($M = 24,20$; $DP = 8,76$) sendo que o valor da Moda apresentou a idade de 18 anos como a mais frequente.

Instrumentos

1. Questionário de Identificação

Este instrumento, construído pelos pesquisadores do Laboratório em Avaliação Psicológica em Saúde Mental III (Lapsam 3) da Universidade São Francisco, foi utilizado com objetivo de obtenção de dados para caracterização da amostra. Ele foi composto por questões que abarcavam informações sobre estado civil, curso (no caso dos estudantes universitários), sexo e idade.

2. Escala Baptista de Depressão – Versão Adulto (EBADEP-A)

A EBADEP-A (Baptista, 2012) é composta por 45 itens, cada qual formado por sentenças que dizem respeito a um descritor sintomático da depressão. Cada item apresenta um par de frases, uma de cunho positivo e outra de cunho negativo, são separadas por uma régua contendo quatro círculos, formando uma escala *Likert* de quatro pontos. Ao participante, é requisitado que ele escolha um dos círculos que melhor represente sua concordância com a frase positiva ou negativa apresentada. Sua pontuação mínima é zero e máxima, 135.

As qualidades psicométricas do instrumento foram aferidas com base na TRI. Os estudos envolveram um banco de dados, construído a partir da coleta de diferentes pesquisas, em vários estados brasileiros, ao longo de aproximadamente três anos (Baptista, 2012). Com base no modelo de Rasch, foi verificada a fidedignidade das estimativas de *Theta* dos respondentes, calculada pelo modelo e considerada a real (igual a 0,92), e no índice dos escores modelados (igual a 0,94), indicando uma quantidade de erro aceitável na avaliação do construto. No que se refere à precisão, foi realizada análise comparativa entre a TCT e a TRI, sendo que tanto o alfa de Cronbach ($\alpha = 0,94$) quanto a precisão real encontrada de acordo com o modelo de Rasch ($\alpha = 0,92$), permitiram a conclusão de que a EBADEP-A possui índices altos de precisão, mais que satisfatórios (Baptista, 2012).

Foram verificadas também, com base em um banco de dados proveniente de aplicações em instituições de ensino superior e da saúde, evidências de validade de construto e critério, além de índices de precisão, tanto pela TCT quanto pela TRI, os quais foram considerados altos e muito satisfatórios, assim como apresentados no manual técnico da escala. No que concerne às evidências de validade de critério, com base na TCT, a

análise de variância determinou a discriminação de cada um dos grupos pertencentes à amostra (Baptista & Gomes, 2011).

3. *Center for Epidemiologic Studies - Depression (CES-D)*

A escala de rastreamento CES-D (Radloff, 1977) é composta por 20 itens, organizados em uma escala *Likert* de quatro pontos, variando de zero a três. Dessa forma o valor mínimo possível é zero e o máximo, 60 pontos, sendo que quanto maior a pontuação, maior a presença de sintomas depressivos. Para o contexto brasileiro, podem ser encontrados estudos de validação da escala em idosos (Batistoni, Neri & Cupertino, 2007; 2010), em populações clínicas e não clínicas (Silveira & Jorge, 1998) e em estudantes universitários (Hauck Filho & Teixeira, 2011; Silveira & Jorge, 1998), sendo que para esta população foi estabelecido um ponto de corte igual a 15 (Silveira & Jorge, 1998). No estudo de Hauck Filho e Teixeira (2011) foram encontrados quatro fatores para a escala, a saber, Depressão (7 itens, $\alpha = 0,86$), Problemas Interpessoais (2 itens $\alpha = 0,70$), Afetos positivos (4 itens, $\alpha = 0,65$) e Aspectos Somáticos (7 itens, $\alpha = 0,75$).

Procedimentos

Após aprovação do Comitê de Ética em Pesquisa da Universidade São Francisco (Protocolo CAE 0422.0.142.000-11), as instituições que previamente autorizaram a realização da pesquisa foram contatadas e os instrumentos foram aplicados em grupos de 20 a 40 pessoas por vez. Os objetivos do estudo foram apresentados aos participantes, os quais assinaram as duas vias do Termo de Consentimento Livre e Esclarecido, e em seguida, os demais instrumentos. Metade dos protocolos foi organizada na seguinte ordem:

questionário de identificação, EBADEP-A, CES-D. A outra parte tal como: questionário de identificação, CES-D, EBADEP-A. O tempo médio de aplicação dos instrumentos foi de 35 minutos.

Resultados

Inicialmente realizou-se procedimento de substituição dos casos *missing* pelas médias respectivas individuais, ressaltando que somente foram aceitos os casos em que mais de 80% dos testes foram respondidos. Por este motivo, foram substituídas no máximo três respostas em cada teste. A EBADEP-A recebeu pontuação mínima de zero e máxima de 116 ($M = 34,02$; $DP = 21,56$), mediana 30 e moda 19, enquanto a CES-D variou entre zero e 47 pontos ($M = 17,28$; $DP = 7,82$), com mediana e moda iguais a 16.

A maioria dos participantes ($N = 510$; 87,8%) foi classificada pela EBADEP-A com ausência de sintomatologia depressiva, de acordo com Baptista (2012), enquanto 45 (7,7%) receberam classificação Leve, 25 (4,3%) Moderada e um participante (0,2%) foi considerado com Sintomatologia Severa, ou seja, com pontuação bruta acima de 111. Quanto à CES-D, foi considerado o ponto de corte 15 apresentado por Silveira e Jorge (1998) como o mais adequado para a amostra universitária. Do total, 316 (54,4%) ficaram acima do ponto de corte, havendo indícios de sintomatologia depressiva.

As análises de correlação de Pearson resultaram em índices de magnitude significativamente alta e positiva entre a EBADEP-A e a CES-D ($r = 0,70$; $p \leq 0,001$). No que se refere às análises entre escores da EBADEP-A e cada um dos fatores da CES-D, os índices foram significativos com magnitude alta em relação ao fator Depressão ($r = 0,75$;

$p \leq 0,001$) e Aspectos Somáticos ($r = 0,73$; $p \leq 0,001$) e de magnitude moderada entre Problemas interpessoais ($r = 0,50$; $p \leq 0,001$) e Afetos positivos ($r = -0,45$; $p \leq 0,001$).

A comparação entre as categorias de gravidade da EBADEP-A em relação à CES-D foi realizada por meio da análise de variância (ANOVA), tendo sido encontrados dados altamente significativos ($F[2, 577] = 149,77$; $p \leq 0,001$). O teste *post hoc* de Tukey revelou três grupos distintos nos quais os participantes classificados com sintomatologia mínima ou ausência de sintomatologia depressiva pela EBADEP-A apresentaram médias significativamente menores ($M = 15,61$; $DP = 6,247$) que de sintomatologia leve ($M = 26,64$; $DP = 6,620$), os quais também apresentaram médias significativamente menores com sintomatologia moderada ($M = 33,48$, $DP = 6,995$). Ressalta-se que para esta análise, o participante o qual recebeu classificação de sintomatologia severa foi retirado para não haver interferência na análise.

O oposto também foi realizado, ou seja, a classificação da CES-D, obtida por meio de ponto de corte, foi comparada aos escores da EBADEP-A. Os resultados encontrados foram altamente significativos ($t[579] = -15,113$; $p \leq 0,001$), e os participantes não depressivos conforme a CES-D ($M = 21,51$; $DP = 12,538$) apresentaram menor média, de forma significativa, que o grupo de participantes com indícios de depressão ($M = 44,52$; $DP = 21,953$).

Foram também investigadas as diferenças em relação ao gênero. O teste *t de Student* revelou dados significativos em relação ao sexo, sendo que os homens obtiveram médias significativamente menores que as mulheres, tanto em relação à EBADEP-A ($t[579] = -0,518$; $p \leq 0,05$) quanto em relação à CES-D ($t[579] = -3,200$; $p \leq 0,001$). Considerando este resultado, foram refeitas as análises de correlação entre os escores brutos das escalas, mas

com controle de sexo, sendo que as magnitudes encontradas nas correlações parciais ($r = 0,69$; $p \leq 0,001$) indicaram a ausência de influência desta variável nas correlações anteriores.

Sobre este ponto específico, dando continuidade aos objetivos do presente estudo, foram realizadas análises com base na Teoria da Resposta ao Item, especificamente pelo modelo de Rasch. Anteriormente ao procedimento de análise de DIF propriamente dito, foi necessário realizar uma análise descritiva da estrutura interna da escala. Inicialmente, no que diz respeito à adequação dos dados ao modelo, foi possível observar que à medida que as opções de resposta aumentam de valor, de zero a três, são acrescidas também as médias esperadas e observadas dos níveis de aderência dos participantes aos itens respondidos (*Theta*), variando entre -2,11 a 0,37 (Tabela 1).

Inserir tabela 1 aproximadamente aqui

Embora a distância entre os limiares observados não tenham sido equivalentes, houve também acréscimo estes valores na medida em que as opções de resposta aumentam (Tabela 1). Além disso, os valores médios esperados não ficaram distantes do que foi observado, o que permite sugerir que no geral a maioria foi respondida dentro de um padrão, e que as médias dos *infits* e *outfits* ficaram próximas a 1,0, não havendo nenhuma média superior a 1,5 mostrando boa adequação ao modelo (Linacre, 2002; 2010).

Para a realização das análises com base no modelo de Rasch, deve ser também considerada a distribuição das opções de resposta dos instrumentos (Linacre, 2002). A Figura 1 apresenta a visualização gráfica da dessa distribuição por meio de curvas equivalentes a cada opção de resposta, em relação às medidas dos itens.

Inserir Figura 1 aproximadamente aqui

Pode-se perceber, pela Figura 1, que embora os limiares (*thresholds*) não estejam com distâncias equivalentes, cada uma das opções de resposta apresentou uma curva característica, sem sobreposição de nenhum limiar. Assim, pode-se observar que a EBADEP-A apresentou uma boa distribuição, com áreas e limiares demarcados. Por sua vez, no que diz respeito aos parâmetros de ajuste dos dados ao modelo, ressalta-se no que diz respeito à análise de ajuste das pessoas, foram desconsiderados os casos extremos, ou seja, iguais a zero ou iguais ao total máximo da escala. No que concerne aos itens, pode-se perceber que as médias dos *infits* e *outfits* foram próximos a 1,0 (Tabela 2).

Inserir Tabela 2 aproximadamente aqui.

Examinando-se de forma detalhada os valores dos *infits* e *outfits* visualizados na Tabela 2, observa-se que a maioria se enquadrou abaixo do valor 1,30 o que pode ser considerado bom sinal de desajuste (Rueda, 2007). Ainda, pode-se considerar um bom sinal que não tenham sido encontrados valores acima de 2, ou seja, com alta incidência de ruídos.

Linacre (2002; 2010) considera que as medidas individuais de *infit* e *outfit* situadas entre 1,3 e 1,5 como ajustadas, porém com um número maior de resíduos, mas não de forma a prejudicar a adequação ao modelo. Foram observados neste caso, os itens relativos à esquivia de situações sociais, desesperança e alterações no sono e libido, no que diz respeito aos *infits* e humor deprimido, alteração no sono, hipocondria, esquivia de situações sociais, e sentimento de incapacidade para os *outfits*. os mesmos três itens relativos à alteração de peso, libido e apetite apresentaram índices entre 1,5 e 2,0 tanto no *infit*,

relativo ao desempenho próximo as habilidades dos sujeitos, quanto no *outfit*, mais distantes dos níveis de habilidade.

Sobre a distribuição dos itens da escala, o que pode ser visualizado no Mapa de Itens (Tabela 3), a variação de dificuldade foi entre -1,04 (Item14) e 1,71 (Item29), sendo que não somente o item 29 apresentou alta dificuldade, mas também o item 44. Uma vez que a EBADEP-A não possui opções de resposta do tipo acerto-e-erro, ou seja, diretamente relacionadas às habilidades dos participantes, mas refere-se à frequência de respostas, pode-se dizer que os referidos aos indicadores de ideação suicida e hipocondria apresentaram baixa frequência de resposta na quarta opção, imediatamente mais próxima à frase de cunho negativo.

Inserir Tabela 3 aproximadamente aqui.

As letras M situadas em ambas os lados da coluna central indicam posição média de um ponto para outro (Tabela 3). Percebe-se que as médias para pessoas e itens foram distintas e que a maior parte dos itens se concentrou entre os valores -1 e 1. Faz-se importante observar também que todos os itens do instrumento se situaram acima da pontuação média dos participantes.

No que se refere ao funcionamento diferencial do item (DIF), diferentes critérios e métodos para a sua detecção podem ser encontrados na literatura (Elosua & López, 2007; Fidalgo & Madeira, 2008; Fidalgo & Ferreres, 2002; Penfield & Lam, 2000). Contudo, de acordo com Linacre (2010), um dos dados mais importantes ao se verificar o DIF é o contraste.

O contraste diz respeito à diferença na dificuldade de cada item entre os dois grupos, o qual apresentar minimamente o valor 0,5 para que o DIF seja percebido (Linacre, 2010). Entretanto, devem-se perceber quais os valores são significantes, e por esta razão, observa-se a probabilidade do valor do contraste em ser ao acaso. Uma vez que para as escalas politômicas a probabilidade Mantel-Haenszel fornece a significância do DIF, este valor deve ser menor que 0,05. A Tabela 4 apresenta as medidas de DIF por Sexo, com seus respectivos desvios padrão, e os valores de contraste e probabilidade medida pelo método Mantel-Haenszel (MH).

Inserir Tabela 4 aproximadamente aqui

A partir destas considerações, foi possível perceber a existência de funcionamento diferencial em sete itens, todos significantes. Os itens 1 e 6 (relacionados ao choro) e o item 42 (sobre desejo sexual), apresentaram funcionamento enviesado ao sexo masculino, enquanto os itens 10 (sobre futuro), 20 (ajudar os outros), 31 (sono) e 32 (realização de tarefas) tiveram funcionamento diferencial enviesado ao grupo feminino.

Discussão

A presente pesquisa realizou uma série de análises com objetivo de buscar evidências de validade para a Escala Baptista de Depressão – versão Adulto (EBADEP-A), com base em uma amostra de estudantes. A maior parte da amostra apresentou sintomatologia mínima, ou ausência de sintomatologia depressiva, constatada pela EBADEP-A (Baptista, 2012), o que é esperado em uma amostra não clínica, entretanto, a maioria ultrapassou o ponto de corte para universitários definido por Silveira e Jorge (1998)

para a CES-D. Tal resultado pode ser explicado pela diferença na constituição das escalas, uma vez que ao se analisar semanticamente as questões da CES-D e dividi-las em grupos de sintomas, pode-se perceber que o instrumento possui foco em sintomas de humor deprimido e fatores afetivos da depressão, enquanto a EBADEP-A possui, em sua estrutura de itens, uma distribuição maior de sintomas somáticos, afetivos, de humor e cognitivos (Baptista, 2012; Radloff, 1977).

Foi possível aferir para a EBADEP-A diferentes evidências de validade baseada na relação com variáveis externas, tanto no que diz respeito aos padrões encontrados nas correlações, com e sem controle do efeito do sexo, quanto no resultado das diferenças de médias averiguadas entre as categorias de classificação da EBADEP-A e os valores do CES-D e as diferenças dos valores médios dos grupos depressivo e não-depressivo, pela CES-D em relação às pontuações na EBADEP-A (AERA, APA & NCME, 1999; Urbina, 2004).

Foram também verificadas as diferenças no que diz respeito aos valores médios obtidos na escala por homens e mulheres, e o presente estudo acompanhou os resultados encontrados em diferentes pesquisas, nas quais as mulheres apresentaram significativamente maiores médias nos instrumentos de depressão (Angst, & cols., 2002; Aros & Yoshida, 2009; Baptista, 2004; Carrillo, Rojo, & Staats, 2004; Justo & Calil, 2006; Kessler, 2003; Nolen-Hoeksema, Larson, & Graysib, 1999; Sprock & Yoder, 1997).

Diversos estudos sobre as possíveis causas para as diferenças podem ser encontrados, desde fatores biológicos até culturais e sociais (Angst, & cols., 2002; Carrillo, Rojo, & Staats, 2004; Kessler, 2003; NIMH, 2000; Piccinelli & Wilkinson, 2000). Entretanto, questiona-se se as diferenças no conteúdo dos instrumentos de avaliação podem

auxiliar ou comprometer para que um grupo de participantes possa obter pontuações maiores ou menores nas escalas. Assim, fala-se do estudo dos possíveis vieses de respostas dos participantes de um grupo ou outro em relação aos itens do instrumento, o qual é verificado pelo funcionamento diferencial do item, por meio da TRI (Elosua & López, 2000; Embrestson & Reise, 2000; Penfield & Lam).

A partir das análises descritivas antecedentes ao procedimento principal de DIF, foi possível perceber a adequação dos dados ao modelo. No que diz respeito à distribuição dos itens em relação aos níveis de habilidade das pessoas, observou-se que todos os itens do instrumento se situaram acima da média das pessoas, o que sugere que a maioria dos participantes apresentou baixos escores no instrumento, assinalando preferencialmente as opções de resposta mais baixas, o que de fato é uma característica comum em amostras de estudantes (Baptista, 2012). Os itens com menor aderência foram os relativos a hipocondria e ideação suicida, o que também é esperado em uma amostra não-clínica.

Especificamente em relação ao funcionamento diferencial, mesmo tendo sido encontrados tanto itens com favorecimento ao grupo masculino quanto ao feminino, quando comparados ao número total de itens da escala, percebe-se que poucos foram o que apresentaram viés a um dos grupos. Por essa razão, pode-se concluir, por um princípio de equidade, que foi observado um equilíbrio entre os grupos em relação aos vieses (Elosua & López, 2000; Embrestson & Reise, 2000; Penfield & Lam ; Rueda, 2007).

Isto posto, pode-se observar que os resultados encontrados relativos às diferenças nas pontuações médias entre os gêneros foi real, e não derivada da forma como o construto é analisado pelos itens da EBADEP-A. O presente estudo apresentou importantes evidências de validade e de consistência interna ao instrumento, tanto pela Teoria Clássica

dos Testes quanto pela Teoria de Resposta ao Item. Entretanto, uma importante limitação ao estudo deve ser ressaltada, isto é, a utilização de uma amostra não clínica na investigação, o que pode fazer com que as pontuações nas escalas fossem mais baixas, com a maior parte com sintomatologia leve ou ausente. Dessa forma, sugere-se a realização de estudos futuros com amostras clínicas, ou seja, compostas por pessoas com diagnóstico confirmado de depressão, a partir dos quais poderão ser investigadas novas evidências de validade para o instrumento.

Referências

- Amarnani, R. (2009). Two Theories, One Theta: A Gentle Introduction to Item Response Theory as an Alternative to Classical Test Theory. *The International Journal of Educational and Psychological Assessment*, 3, 104-109.
- American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME]. (1999). *Standards for Psychology and Educational Testing*. Washington, DC: American Psychology Association.
- American Psychiatric Association [APA]. (2002). *DSM-IV-TR - Manual diagnóstico e estatístico de transtornos mentais*. (4ª ed). Porto Alegre: Artmed.
- Angst, J., Gamma, A., Gastpar, M., Lépine, J-P., Mendlewicz, J., & Tylee, A. (2002). Gender differences in depression. *European Archives of Psychiatry and Clinical Neuroscience*, 252, 201-209.
- Aros, M. S., & Yoshida, E. M. P. (2009). Estudos da depressão: instrumentos de avaliação e gênero. *Boletim de Psicologia*, 59(130), 61-76.
- Baptista, M. N. (2004). *Suicídio e Depressão: Atualizações*. São Paulo: Guanabara Koogan.
- Baptista, M. N. (2012). *Manual Técnico da Escala Baptista de Depressão – Versão Adulto (EBADEP-A)*. São Paulo: Vetor.
- Baptista, M. N., & Carneiro, A. M. (2011). Validade da escala de depressão: relação com ansiedade e *stress* laboral. *Estudos de Psicologia*, 28(3), 345-352.
- Baptista, M. N., & Gomes, J. O. (2011). Escala Baptista de Depressão (Versão Adulto) – EBADEP-A: evidências de validade de construto e de critério. *Psico-USF*, 16(2), 151-161.

- Baptista, M. N., Argimon, I. I. L., & Yoshida, E. M. P. (2012). Escalas de Avaliação de Sintomas Depressivos. Em E. Boruchovitch, A. A. A. Santos, & E. Nascimento. *Avaliação Psicológica nos Contextos Educativo e Psicossocial*. São Paulo: Casa do Psicólogo.
- Baptista, M. N., Cardoso, H. F., & Gomes, J. O. (2012). Escala Baptista de Depressão (Versão Adulto) - EBADEP-A: validade convergente e estabilidade temporal. *Psico-USF*, 17(3).
- Baptista, M. N., Souza, M. S., & Alves, G. A. (2008). Evidências de validade entre a Escala de Depressão (EDEP), o BDI e o Inventário de Percepção de Suporte Familiar (IPSF). *Psico-USF*, 13(2), 211-220.
- Batistoni, S. S. T., Néri, A. N., & Cupertino, A. P. (2010). Validade e confiabilidade da versão Brasileira da Center for Epidemiological Scale – Depression (CES-D) em idosos Brasileiros. *Psico-USF*, 15(1), 13-22.
- Batistoni, S. S. T., Néri, A. N., & Cupertino, A. P. F. B. (2007). Validade da escala de depressão do Center for Epidemiological Studies entre idosos Brasileiros. *Revista de Saúde Pública*, 41(4), 598-605.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1997). *Terapia da Depressão*. Rio de Janeiro: Zahar.
- Calil, H. M., & Pires, M. L. N. (1998). Aspectos gerais das escalas de avaliação de depressão. *Revista de Psiquiatria Clínica*, 25(5), 240-244.
- Carrillo, J. M., Rojo, N., & Staats, A. W. (2004). Women and vulnerability to depression: some personality and clinical factors. *The Spanish Journal of Psychology*, 7(1), 29-39.

- Conselho Federal de Psicologia – CFP (2003). Resolução nº 002/2003 [On-line]. Recuperado de http://www.pol.org.br/pol/export/sites/default/pol/legislacao/legislacaoDocumentos/resolucao2003_02.pdf
- Coutinho, M. P. L., Gontiès, B., Araújo, L. F., & Sá, R. C. N. (2003). Depressão, um sofrimento sem fronteira: representações sociais entre crianças e idosos. *Psico-USF*, 8(2), 183-192.
- Cronbach L. J. (1996). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artes Médicas.
- Elosua, P., & López, A. J. (2007). Aplicación de cuatro procedimientos de detección del funcionamiento diferencial sobre ítems politómicos. *Psicothema*, 19(2), 329-336.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum.
- Ferster, C. B., Culbertson, S., & Boren, C. P. (1977). *Princípios do comportamento*. (trad: Maria Ignez Rocha e Silva, Maria Alice de Campos Rodrigues e Maria Benedita Lima Pardo). São Paulo: Hucitec.
- Fidalgo, A. M. A., & Ferreres, D. T. (2002). Supuestos y consideraciones en los estudios empíricos sobre el funcionamiento diferencial de los ítems. *Psicothema*, 14(2), 491-496.
- Fidalgo, A. M., & Madeira, J. M. (2008). Generalized Mantel-Haenszel Methods for Differential Item Functioning Detection. *Educational and Psychological Measurement*, 68(6), 940-958.

- Fleck, M. P. A., Lafer, B., Sougey, E. B., Del Porto, J. A., Brasil, M. A., & Jurena, M. F. (2003). Diretrizes da Associação Médica Brasileira para o tratamento da depressão (versão integral). *Revista Brasileira de Psiquiatria*, 25(2), 114-122.
- Gomes, J. O., & Baptista, M. N. (2010). Escala de Depressão (EDEP) e medidas de atenção dividida e sustentada em universitários. *Boletim de Psicologia*, 60(133), 191-204.
- Gorenstein, C., & Andrade, L. H. S. G. (1998). Inventário de Depressão de Beck: propriedades psicométricas da versão em português. *Revista de Psiquiatria Clínica*, 25(5), 245-250.
- Hambleton, R. K., & Jones, R. W. (1993). An NCME Instructional Module on Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Educational Measurement: Issues and Practice*, 12(3), 38-47.
- Hauck Filho, N., & Teixeira, M. A. P. (2011). A estrutura fatorial da escala CES-D em estudantes universitários brasileiros. *Avaliação Psicológica*, 10(1), 91-97.
- Justo, L. P., & Calil, H. M. (2006). Depressão – o mesmo acometimento para homens e mulheres? *Revista de Psiquiatria Clínica*, 33(2), 74-79.
- Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders*, 74, 5-13.
- Linacre, J.M. (2010). Winsteps® (Version 3.70.0.2) [Computer Software]. Beaverton, Oregon: Winsteps.com.
- Linacre J. M. (2002). What do Infit and Outfit, Mean-Squared and Standardized mean? *Rasch Measurement Transactions*, 16 (2), 878. Recuperado de <http://209.238.26.90/rmt/rmt82a.htm>

- Marcolino, J. A. M., Mathias, L. A. S. T., Piccinini-Filho, L., Guaratini, A. A., Suzuki, F. M., & Alli, L. A. C. (2007). Escala hospitalar de ansiedade e depressão: estudo da validade de critério e da confiabilidade com pacientes no pré-operatório. *Revista Brasileira de Anestesiologia*, 57(1), 52-62.
- Matos, E. G., Matos, T. M. G., & Matos, G. M. G. (2006). Depressão Melancólica e Depressão Atípica: aspectos clínicos e psicodinâmicos. *Estudos de Psicologia*, 23(2), 173-179.
- Moreno, R. A., & Moreno, D. H. (1998). Escalas de depressão de Montgomery & Åsberg (MADRS) e de Hamilton (HAM-D). *Revista de Psiquiatria Clínica*, 25(5), 1-17.
- National Institute of Mental Health (NIMH) (2000). *Depression Research*. Office of communications and Public Liaison National Institute of Mental Health. Bethesda, MD.
- Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the Gender Difference in Depressive Symptoms. *Journal of Personality and Social Psychology*, 77(5), 1061-1072.
- Noronha, A. P. P., & Alchieri, J. C. (2002). Reflexões sobre os Instrumentos de Avaliação Psicológica. Em Ricardo Primi (org.). *Temas em avaliação psicológica*. Campinas: Imprensa Digital do Brasil
- Noronha, A. P. P., & Alchieri, J. C. (2004). Conhecimento em Avaliação Psicológica. *Estudos de Psicologia*, 21(1), 43-52.
- Noronha, A. P. P., & Vendramini, C. M. M. (2003). Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica*, 16(1), 177-182.

- Nunes, C. H. S. S., & Primi, R. (2009). Teoria de resposta ao item: conceitos e aplicações na psicologia e na educação. Em Hutz, C. S. (org). *Avanços e polêmicas em avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Organização Mundial da Saúde – OMS (2001). *Relatório Mundial de Saúde*. Saúde Mental: nova concepção, nova esperança. [On-Line]. Recuperado de http://www.who.int/whr/2001/en/whr01_djmessage_po.pdf
- Organização Mundial da Saúde (OMS) (1993). *Classificação dos Transtornos Mentais e do comportamento - CID-10: descrições e diretrizes diagnósticas*. Trad. Dorgival Caetano. (3º Volume, 10ª Ed.). Porto Alegre: Artes Médicas.
- Pasquali, L. (1998). *Psicometria: Teoria e aplicações*. Brasília: Editora UnB.
- Pasquali, L., & Primi, R. (2003). Fundamentos da Teoria da Resposta ao Item – TRI. *Avaliação Psicológica*, 2(2), 99-110.
- Pasquali, L. (1999). *Instrumentos psicológicos: manual prático de elaboração*. Brasília: LabPAM & IBAPP.
- Penfield, R. D., & Lam, T. C. M. (2000). Assessing Differential Item Functioning in Performance Assessment: review and recommendations. *Educational Measurement: Issues and Practice*, 19(3), 5-15.
- Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression: critical review. *British Journal of Psychiatry*, 177, 486-492.
- Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, 1(3), 385-401.
- Rueda, F. J. M. (2007). O funcionamento diferencial do item no teste pictórico de memória. *Revista Avaliação Psicológica*, 6(2), 229-237.

- Santor, D. A., Gregus, M., & Welch, A. (2006). Eight Decades of Measurement in Depression. *Measurement*, 4(3), 135-155.
- Santos, I. S., Matijasevich, A., Tavares, B. F., Barros, A. J. D., Botelho, I. P., Lapolli, C., Magalhães, P. V. S., Barbosa, A. P. P. N., & Barros, F. C.. (2007). Validation of the Edinburgh Postnatal Depression Scale (EPDS) in a sample of mothers from the 2004 Pelotas Birth Cohort Study. *Cadernos de Saúde Pública*, 23(11), 2577-2588.
- Silveira, D. X., & Jorge, M. R. (1998). Propriedades psicométricas da escala de rastreamento populacional para depressão CES-D em populações clínica e não-clínica de adolescentes e adultos jovens. *Revista de Psiquiatria Clínica*, 25(5), 251-61.
- Sprock, J., & Yoder, C. Y. (1997). Women and Depression: an update on the report of the APA Task Force. *Sex Roles*, 36(516), 269-303.
- Suija, K., Kalda, R., & Maaros, H-I. (2009). Patients with depressive disorder, their co-morbidity, visiting rate and disability in relation to self-evaluation on physical and mental health: a cross-sectional study in family practice. *BMC Family Practice*, 10 (38).
- Urbina, S. (2004). *Essentials of Psychological Testing*. New Jersey: John Wiley & Sons, Inc.
- Valentini, F., & Laros, J. A. (2011). Teoria de Resposta ao Item na Avaliação Psicologia. Em R. A. M. Ambiel, I. S. Rabelo, S. V. Pacanaro, G. A. S. Alves, I. F. A. S. Leme (Orgs). *Avaliação Psicológica: guia de consulta para estudantes e profissionais de psicologia*. São Paulo: Casa do Psicólogo.

- Vicente, B., Rioseco, P., Saldivia, S., Kohn, R., & Torres, S. (2005). Prevalencia de trastornos psiquiátricos en Latinoamérica: revisión crítica. *Revista Colombiana de Psiquiatría*, 34(4), 506-514.
- Vismari, L. (2004). Neuroanatomia dos transtornos de humor. *ConScientiae Saúde*, 3, 95-102.

Tabela 1

Sumário de informações sobre a EBADEP-A de acordo com o modelo de Rasch

| Opções de resposta | Escores observados | % | Média Observada | Média Esperada | <i>Infit</i> | <i>Outfit</i> | Limiares |
|--------------------|--------------------|----|-----------------|----------------|--------------|---------------|----------|
| 0 | 12412 | 47 | -2,11 | -2,09 | 1,02 | 1,02 | |
| 1 | 9111 | 35 | -1,07 | -1,09 | 0,93 | 0,89 | -1,27 |
| 2 | 3209 | 12 | -0,24 | -0,28 | 0,94 | 0,95 | 0,37 |
| 3 | 1413 | 5 | 0,37 | 0,46 | 1,11 | 1,32 | 0,91 |

Tabela 2

Parâmetros de ajuste dos itens e das pessoas, com exclusão de casos extremos

| Parâmetros | Itens (N=45) | | | | Pessoas (N=577) | | | |
|-----------------|-----------------|------------|-------|------|--------------------|--------|-------|------|
| | Infit | Outfit | b | erro | Infit | Outfit | Theta | erro |
| Média | 1,02 | 1,01 | 0,00 | 0,06 | 1,01 | 1,01 | -1,38 | 0,25 |
| DP | 0,25 | 0,27 | 0,67 | 0,01 | 0,41 | 0,42 | 1,08 | 0,09 |
| Máximo | 1,62 | 1,66 | 1,71 | 0,10 | 3,16 | 4,84 | 1,94 | 1,01 |
| Mínimo | 0,59 | 0,60 | -1,04 | 0,05 | 0,28 | 0,29 | -5,25 | 0,18 |
| Menor que 1,3 | 38 (84,44%) | 36 (80%) | | | | | | |
| Entre 1,3 e 1,5 | 4 (8,89%) | 6 (13,33%) | | | | | | |
| Entre 1,5 e 2,0 | 3 (6,67%) | 3 (6,67%) | | | | | | |

Tabela 4

Itens com funcionamento diferencial por sexo, de acordo com a medida de contraste

| Sexo | <i>b</i> | DP | Sexo | <i>b</i> | DP | Contraste | DP | <i>p</i> (MH) | Itens |
|-----------|----------|------|----------|----------|------|-----------|------|---------------|--------|
| Masculino | -0,17 | 0,12 | Feminino | -1,03 | 0,06 | 0,86 | 0,14 | 0 | Item1 |
| Masculino | 0,69 | 0,14 | Feminino | -0,67 | 0,06 | 1,35 | 0,16 | 0 | Item6 |
| Masculino | 0,61 | 0,15 | Feminino | 1,21 | 0,1 | -0,6 | 0,18 | 0,0035 | Item10 |
| Masculino | -0,29 | 0,12 | Feminino | 0,41 | 0,08 | -0,7 | 0,14 | 0,0002 | Item20 |
| Masculino | -0,51 | 0,12 | Feminino | 0,04 | 0,07 | -0,55 | 0,14 | 0,0026 | Item31 |
| Masculino | -0,02 | 0,13 | Feminino | 0,54 | 0,08 | -0,56 | 0,15 | 0,0002 | Item32 |
| Masculino | 1,12 | 0,17 | Feminino | 0,04 | 0,07 | 1,08 | 0,18 | 0 | Item42 |

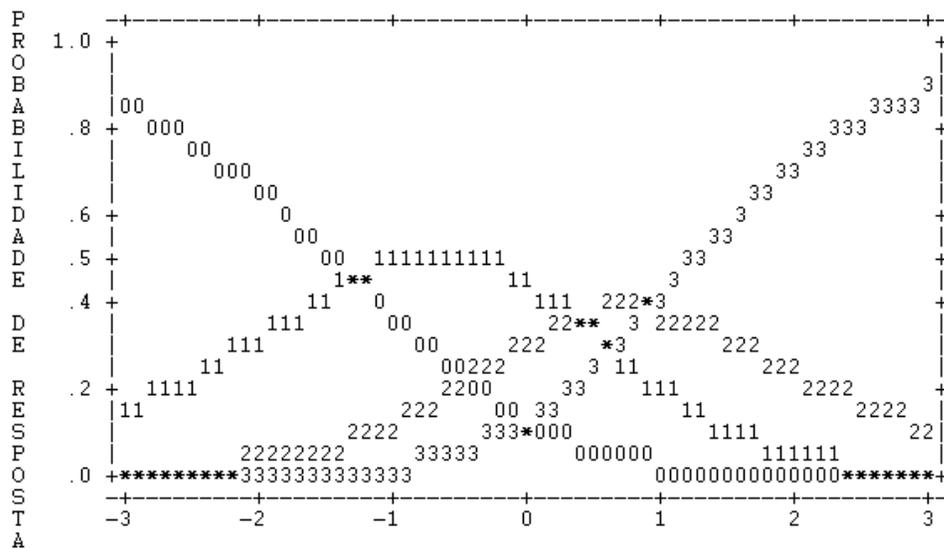


Figura 1. Gráfico de probabilidades de resposta em relação a cada opção de resposta da EBADEP-A

Artigo 2

Análise Fatorial confirmatória da Escala Baptista de Depressão – Versão

Adulto

A depressão é um dos construtos mais estudados mundialmente e são diversos os meios de se mensurá-la, por diferentes métodos. No Brasil, podem ser encontrados vários instrumentos validados para diversas amostras, mas somente um instrumento construído no país, a Escala Baptista de Depressão – Versão Adulto (EBADEP-A). Ao longo de sua construção, foram publicados estudos de evidências de validade e precisão, somente duas organizações fatoriais foram apresentadas. Esta pesquisa teve como objetivo realizar um estudo fatorial confirmatório para a EBADEP-A. Participaram da pesquisa 632 estudantes, com média de idade de 23,72 anos, majoritariamente mulheres (74,8%). Foram testados os modelos uni e bifatorial e verificados os índices de ajuste em ambos os modelos, dentro dos quais os valores RMSEA encontrados foram bastante semelhantes. A análise exploratória complementar pode confirmar a estrutura unifatorial como a de melhor ajuste, pelos critérios de *Scree Plot* e porcentagem de variância explicada.

Palavras-chave: análise multivariada; transtorno de humor; avaliação psicológica

Confirmatory Factor Analysis of the Baptista Scale of Depression – Adult Version

Depression is one of the most studied constructs worldwide and there are several ways to measure it, by different methods. In Brazil, it can be found different validated instruments for several samples, but only one instrument built in this country, the Baptista Scale of Depression – Adult Version (EBADEP-A). Throughout its construction different studies about its validity and reliability have been published, but only two factorial organizations were presented. This research aimed to conduct a confirmatory factorial study for the EBADEP-A. Participants were 632 students, mean age of 23.72 years, mostly women (74.8%). Uni and bifactorial models were tested and verified fit indexes in both models, within which the RMSEA values found were very similar. The additional exploratory analysis was able to confirm the one factor structure as with the best fit, by the Scree Plot and percentage of explained variance criteria.

Keywords: Multivariate analysis; Mood Disorder; Psychological Assessment

Análisis factorial confirmatorio de la Escala Baptista de Depresión - Versión Adultos

La depresión es uno de los constructos en más estudiados todo el mundo, y hay varias maneras de medir, por diferentes métodos. En Brasil, se pueden encontrar diferentes instrumentos validados para varias muestras, y sólo un instrumento construido en el país, la Escala Baptista de Depresión - Versión Adultos (EBADEP-A). En su construcción diferentes estudios de validez y confiabilidad se han publicado, pero solo dos organizaciones factoriales se presentaron. Esta investigación tuvo como objetivo realizar un estudio factorial confirmatorio para la EBADEP-A. Los participantes fueron 632 estudiantes, con una edad media de 23,72 años, en su mayoría mujeres (74,8%). Los modelos uni y bifactorial fueron probados y verificados los índices de ajuste para ambos modelos, en el que los valores RMSEA encontrados fueron muy similares. El análisis exploratoria adicional puede confirmar la la estructura con un factor como la mejor opción, como criterios el gráfico de sedimentación y porcentaje de varianza explicada.

Palabras clave: análisis multivariante; trastorno de humor; evaluación psicológica

Introdução

De acordo com a Classificação Internacional de Doenças [CID-10] (OMS, 1993), a depressão é um transtorno afetivo, e está categorizada na sessão “Episódios Depressivos” e também no “Transtorno Depressivo Recorrente”. Então, pode-se entender que para a Organização Mundial de Saúde (OMS), a depressão não é uma doença que se manifesta de maneira contínua, mas que se apresenta na forma de episódios, mesmo sendo considerada como crônica e recorrente (Fleck & cols., 2003). Ainda, para a OMS, uma pessoa pode sofrer mais de um episódio ao longo da vida, e a sua identificação depende da presença tanto de sintomas obrigatórios, quanto de alguns adicionais, que caracterizam o perfil da doença. Entretanto, se um novo episódio ocorre em uma situação que pode se chamar de “remissão”, ele passa a ser classificado como Transtorno Recorrente (OMS, 1993).

Para o Manual Diagnóstico e Estatístico de Transtornos Mentais [DSM-IV-TR] (APA, 2002), a depressão está qualificada como um transtorno de humor dentro das especificações “Transtorno Depressivo Maior” e “Transtorno Distímico”, este último comumente chamado de depressão menor. Diferentemente do manual da OMS, a classificação da *American Psychiatric Association* não apresenta a depressão dentro de episódios separados, mas considera a forma com que os sintomas interferem na vida diária da pessoa.

Para ambos os manuais diagnósticos, são considerados tanto o número de sintomas observados quanto a qualidade e a intensidade com que eles se apresentam. Para o CID-10 (OMS, 1993), a depressão pode receber diferentes classificações, a saber, leve, quando verificados dois ou três sintomas; moderada, ao serem identificados quatro ou mais sintomas; grave sem sintomas psicóticos, quando percebidos mais de seis sintomas,

geralmente acompanhados de ideação suicida; e grave com sintomas psicóticos, quando há mais de seis sintomas, com a presença de alucinações e ideias delirantes. Para o transtorno recorrente, deve-se constatar mais de um episódio depressivo em menos de seis meses.

Por sua vez, para DSM-IV-TR (APA, 2002), somente o Transtorno Depressivo Maior recebe categorizações, quais sejam, leve, moderada, severa com ou sem aspectos psicóticos. Diferentemente do CID-10, no qual se utiliza um *checklist* para a identificação dos episódios (OMS, 1993), para o manual da APA, não é apresentado o número de sintomas mínimos para cada categorização, uma vez que o diagnóstico é realizado por meio de uma entrevista padronizada, denominada *Structured Clinical Interview for the DSM Disorders*, a partir da qual pontua-se a magnitude com que cada sintoma é anunciado, e a forma como eles interferem ou prejudicam nas atividades cotidianas.

Os sintomas obrigatórios para a identificação da doença incluem tristeza, ou humor deprimido, e/ ou a falta de energia ou perda de interesse na realização de atividades prazerosas (anedonia). Dentre os sintomas específicos ou adicionais, podem ser citados alterações no apetite ou peso, sono e atividade psicomotora; cansaço, mesmo após esforço mínimo; diminuição da energia, da autoestima e autoconfiança; sentimentos de desvalia ou culpa e indignidade; dificuldades para pensar, concentrar-se ou tomar decisões, além de pensamentos recorrentes sobre morte ou ideação suicida, planos ou tentativas de suicídio (APA, 2002; OMS, 1993).

No que diz respeito à avaliação da depressão, não somente os manuais diagnósticos internacionais são encontrados, mas diferentes maneiras de se mensurá-la, algumas apresentadas na forma de entrevistas estruturadas, completamente guiadas por um profissional devidamente treinado para sua administração, pontuação e interpretação, ou

instrumentos de autoavaliação, os quais são respondidos diretamente pelo indivíduo, após instruções dadas pelo profissional. Há também os de avaliação global, que envolvem a observação do profissional responsável e os mistos, que abarcam ambos os procedimentos (Calil & Pires, 1998).

Como exemplo de entrevista para o diagnóstico de depressão, podem ser citadas a Escala de Hamilton (HAM-D) e a Entrevista Clínica Estruturada para os Transtornos do Eixo I (SCID-CV, do Inglês *Structured Clinical Interview for the DSM Disorders – Clinic Version*). A administração de entrevistas, tanto estruturadas como semiestruturadas, oferecem por um lado uma avaliação pormenorizada e individual das características e sintomas, o que permite um diagnóstico preciso. Por outro lado, seu uso pode se tornar algumas vezes inviável, por exemplo, caso o treinamento tenha sido mal realizado, ou tenha havido mau entendimento do profissional acerca dos critérios de pontuação, ou ainda em relação ao tempo necessário para se investigar completamente o transtorno (Baptista, 2004; Del-Bem, & cols., 2001; Calil & Pires, 1998).

Já as escalas de autoavaliação, embora não tenham objetivo diagnóstico, são mais econômicas, de aplicação rápida, além de serem de fácil correção por parte dos profissionais. Para sua efetiva aplicação, é necessário um nível mínimo de escolaridade para que os itens possam ser corretamente compreendidos. Além disso, este tipo de testagem não é aconselhável para pacientes com psicopatologia grave, pois são altas as chances de se obter respostas com “efeito teto”, ou seja, uma tendência do paciente a sempre escolher a pior opção, sem analisar de maneira eficaz os itens e as alternativas (Calil & Pires, 1998; Moreno & Moreno, 2005).

Vários são os instrumentos autoaplicáveis reconhecidos pela comunidade científica para a avaliação da depressão, tais como o Inventário de Depressão de Beck (BDI), em suas formas BDI-I e BDI-II, a *Center for Epidemiologic Studies* (CES-D), a Escala de Depressão de Montgomery-Åsberg (MADS) e a escala de autoavaliação para Depressão de Zung. Dentre os instrumentos destinados a grupos específicos, podem ser citados como exemplo a Escala Geriátrica de Depressão (GDS), nas versões com 15 e 30 itens, a Escala Hospitalar de Ansiedade e Depressão (*Hospital Anxiety and Depression* [HAD]), a Escala de Auto-registro de Depressão Pós-Parto (*Edinburgh Postnatal Depression Scale*) e o Inventário de Depressão Infantil (CDI).

Em pesquisa internacional, Santor, Gregus e Welch (2006) identificaram em artigos científicos publicados cerca de 280 medidas para a depressão, 70 no formato de escalas, havendo destaque para o BDI, CES-D, Escala de Hamilton, MADS e o SCL90. No Brasil, em pesquisa sobre os aspectos gerais das escalas de depressão, Calil e Pires (1998) classificaram o BDI e a Escala de Depressão de Zung como as escalas de autoaplicação mais utilizadas e a Escala de Hamilton, como a entrevista estruturada mais escolhida por profissionais para a avaliação da depressão. Por sua vez, Baptista, Argimon e Yoshida (2012) encontraram 30 instrumentos em 200 artigos, publicados em 54 periódicos, sendo o BDI o mais utilizado, seguido pelo CDI, HAD, GDS, Escala de Hamilton e a Escala de Auto-registro de Depressão Pós-Parto.

Interessante ressaltar que, especificamente no que se refere aos sintomas, tanto os priorizados pelos manuais diagnósticos quanto pelos instrumentos de avaliação da depressão, são verificadas características somáticas, como alterações no apetite, peso, sono; aspectos cognitivos, como a dificuldade na concentração, sentimentos de desvalia,

indignidade e culpa; e também aspectos sociais, como a perda de interesse na realização de atividades, por exemplo (Abreu, 2006; APA, 2002; Beck, Rush, Shaw, & Emery, 1997; Cavalcante, 1997; Del Porto, 1999; Matos, Matos & Matos, 2006; OMS, 1993). Tais evidências levaram pesquisadores como Calil e Pires (1998) e Santor, Gregus e Welch (2006) a verificarem os tipos de sintomas mais presentes nas mais variadas escalas de avaliação da depressão publicadas ao longo dos anos.

Calil e Pires (1998) apresentaram um quadro comparativo no qual as diversas escalas de depressão foram organizadas em relação às características sintomáticas de seus itens, de forma a obter quais as prioridades de cada instrumento. No artigo, são citados os seguintes instrumentos: Escala de Avaliação de Depressão de Hamilton; Escala de Avaliação de Depressão de Montgomery-Asberg; Escala de Avaliação de Melancolia de Bech-Rafaelsen; Inventário de Depressão de Beck; Escala de Auto-Avaliação da Depressão de Zung; Inventário de Auto-Avaliação de Wakefield; e Escala de Avaliação da Depressão de Carrol. As categorias de sintomas mais encontradas nas escalas foram os Cognitivos, Vegetativos ou somáticos e de Humor. As demais categorias citadas foram: sintomas Motores, Sociais, de Ansiedade e Irritabilidade.

Santor, Gregus e Welch (2006) analisaram as medidas de depressão em estudos sobre a sua avaliação ao longo de oito décadas, entre 1918 e 2000. Os sintomas mais presentes nas escalas foram humor deprimido, anedonia e desesperança. Ao longo dos anos pôde-se notar um aumento no número de itens que investigam o comportamento e a ideação suicida, ao passo em que foi sendo observada diminuição no número de itens envolvendo ansiedade.

Tais achados podem ser interpretados como uma resposta à forma como a depressão passou a ser estudada e operacionalizada clínica e cientificamente. Por um lado, a ansiedade foi deixando de ser avaliada como uma parte da depressão, recebendo uma classificação específica nos manuais diagnósticos, com sintomas e critérios determinados (APA, 2002; OMS, 1993). Por outro lado, sabe-se que, embora a desesperança e o comportamento suicida não sejam sintomas classificados como obrigatórios para o diagnóstico e avaliação, podem ser encontradas diversas pesquisas envolvendo a relação entre ambas (Au, Lau, & Lee, 2009; Chachamovich, Stefanello, Botega, & Turecki, 2009; Grassi-Oliveira & cols., 2012; Millis & Kroner, 2008; Panzarella, Alloy, & Whitehouse; 2006; Suija, Kalda, & Maaroos, 2009).

No que diz respeito aos instrumentos que apresentam alguma adaptação para a população brasileira, ou estudos psicométricos e de validade, independente da faixa etária, encontram-se a HAM-D, CES-D, CDI, HAD, a Escala de Auto-Registro de Depressão Pós-Parto, a Escala Geriátrica de Depressão, os Inventários de Depressão de Beck, BDI-I e BDI-II, e a Escala Baptista de Depressão – Versão Adulto (EBADEP-A), sendo estes três últimos avaliados e aprovados pelo Conselho Federal de Psicologia (Baptista, 2012; Calil & Pires, 1998; CFP, 2011; Coutinho, Carolino, & Medeiros, 2008; Gorestein & Andrade, 1998; Moreno & Moreno, 2005; Paradela, Lourenço, & Veras, 2005; Santor, Gregus & Welch, 2006; Santos & cols., 2007).

A EBADEP-A foi construída a partir de descritores de depressão, provindos nos manuais diagnósticos internacionais e nas teorias Cognitiva de Beck e Comportamental de Ferster (APA, 2002; Beck, & cols., 1997; Ferster, Culbertson & Boren, 1977; OMS, 1993) e por isso ela possui em sua distribuição de itens que abordam sintomas cognitivos (33%),

de humor (20%), sociais (18%), vegetativos ou somáticos (18%), motores (4,5%), de irritabilidade (4,5%) e ansiedade (2%) (Baptista, 2012). Com essas características, é pertinente questionar, de forma técnica, como se dá a organização destes sintomas presentes nos itens do instrumento.

Em estudos iniciais durante o seu processo de construção, quando ainda era composta por 75 itens, foram realizadas por Souza (2010) um total de cinco Análises de Componentes Principais com objetivo de investigar evidências de validade e precisão para o instrumento. Compuseram a amostra 771 participantes, com média de idade de 26 anos (DP = 8,60), majoritariamente mulheres (70,2%). A melhor solução fatorial apresentada dividiu a estrutura da escala em dois fatores, explicando 35,32% da variância total (23,30% no primeiro fator), considerando a exclusão de 32 itens. O primeiro fator foi denominado afetivo-cognitivo e agrupou itens referentes aos sintomas de humor deprimido, perda ou diminuição de prazer, desesperança, desamparo, autoestima rebaixada, sentimento de incapacidade, inadequação e inutilidade, carência/dependência, negativismo, esquiva de situações sociais, queda de produtividade, pensamento de morte, e falta de perspectiva sobre o futuro. O segundo, nomeado como Somático, congregou predominantemente os itens relativos à alteração de apetite e/ou peso, insônia, lentidão psicomotora e fadiga /perda de energia, além de dificuldade em concentração e irritabilidade.

No que se refere à versão oficial e finalizada do instrumento, composta por 45 itens, Baptista (2012) realizou diferentes análises com objetivo de investigar sua consistência interna, comparando os índices de precisão entre a Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI), sendo que tanto o alfa de Cronbach ($\alpha = 0,94$) quanto a precisão real encontrada de acordo com a TRI ($\alpha = 0,92$), permitiram a conclusão de que a

EBADEP-A possui índices altos, mais que satisfatórios de confiabilidade. Além disso, a partir da análise por componentes principais de resíduos, foram apontados 39,1% de variância explicada pelos itens, apresentando a proposta de unifatorialidade da escala.

Dessa forma, percebe-se que a EBADEP-A, em sua versão final, é apresentada como unifatorial (Baptista, 2012), embora contenha as mesmas categorias sintomáticas que gerou a organização em dois fatores, nos estudos iniciais de Souza (2010). Por essa razão, pode-se levantar o questionamento sobre qual a melhor estrutura para o instrumento em termos de organização fatorial, isto é, se uni ou bifatorial, neste caso, separando-se os descritores somáticos, de irritação e dificuldade de concentração, dos afetivos e cognitivos.

De fato, embora a EBADEP-A, com 45 itens, possa apresentar a mesma divisão categórica no que se refere aos descritores sintomáticos que sua versão inicial, com 75 itens, o modelo unifatorial parece ser o mais adequado, embora não tenham sido encontradas pesquisas que pudessem corroborar ou confrontar esta hipótese. Um meio de se avaliar qual estrutura para os conjuntos de dados apresentados pode ser considerado a mais consistente para representar o construto é a Análise Fatorial Confirmatória [AFC] (Brown, 2006; Cole, 1987; Kline, 2004). Trata-se de um modelo teórico-técnico, o qual permite analisar a estrutura dos instrumentos de pesquisa frente à população estudada (Cole, 1987).

A AFC é um método propício para confrontar a questão levantada na presente pesquisa, pois, diferentemente da análise exploratória, ela permite a definição *a priori* das estruturas fatoriais a serem testadas (Hambleton & Jones, 1993). Por meio deste procedimento estatístico, a mensuração de um construto tão subjetivo como a depressão pode ser investigado em termos matemáticos e gráficos, por meio da visualização da relação entre as variáveis medidas pelo instrumento (Brown, 2006). Assim, o objetivo do

presente estudo foi a realização de um estudo fatorial confirmatório para a Escala Baptista de Depressão – Versão Adulto (EBADEP-A), a fim de que ambas as estruturas teorizadas da escala possam ser comparadas em termos de adequação aos modelos propostos.

Método

Participantes

O presente estudo contou com a colaboração de 632 estudantes, de diferentes regiões interioranas de São Paulo (Lins, Itatiba, Bauru) e Minas Gerais (Juiz de Fora, Viçosa, Caratinga), a maioria do sexo feminino ($n = 473$; 74,8%). Do total 95,4% eram universitários, tanto de cursos de graduação quanto de Pós-graduação *Lato Sensu*, enquanto que os 29 estudantes não universitários (4,6%) faziam parte de um curso profissionalizante de uma das cidades do interior de São Paulo. As idades variaram entre 17 e 63 anos ($M = 23,72$; $DP = 8,64$), sendo que o valor da Moda apresentou a idade de 18 anos como a mais frequente.

Instrumentos

1. Questionário de Identificação

O objetivo deste questionário, elaborado por estudantes e professores do Laboratório em Avaliação Psicológica em Saúde Mental III (Lapsam 3) da Universidade São Francisco, foi coletar informações para a caracterização da amostra, como cidade, sexo e idade.

2. Escala Baptista de Depressão (EBABEDP-A)

Construída por Baptista (2012), esta escala tem como objetivo avaliar a presença de sintomatologia depressiva em adultos, com base na investigação de questões com cunho positivo e negativo em relação a 26 descritores de depressão. Ela é composta por 45 itens, apresentados em pares de frases. Cada par é separado por uma régua contendo quatro círculos, os quais são assinalados de forma a mostrar a forma como a pessoa concorda com o descritor apresentado, o que caracteriza a escala como sendo do tipo *Likert*, de quatro pontos. Ao participante é requisitado que seja realizada leitura de ambas as frases para a escolha das opções de resposta. O instrumento possui diferentes estudos com evidências de validade e confiabilidade, tanto pela Teoria Clássica dos Testes quanto pela Teoria da Resposta ao Item (Baptista, 2012; Baptista & Gomes, 2011; Baptista, Cardoso, & Gomes, 2012), além de ser um instrumento com parecer favorável pelo Sistema de Avaliação de Testes Psicológicos (CFP, 2011).

Procedimentos

Após contato inicial com as instituições de ensino para autorização à realização do presente estudo, buscou-se aprovação do Comitê de Ética em Pesquisa da Universidade São Francisco (Protocolo CAE 0422.0.142.000-11), após a qual foram agendados os procedimentos de coletas de dados. Os instrumentos foram aplicados em grupos, de 20 a 40 alunos por vez, tendo sido apresentados inicialmente os objetivos da pesquisa para assinatura (em duas vias) do Termo de Consentimento Livre e Esclarecido. Como procedimentos finais, foram aplicados os critérios de exclusão de protocolos para análises,

a partir dos quais foram desconsiderados os protocolos com mais de duas questões em branco na EBADEP-A e os respondidos por estudantes menores de 17 anos, aceitos na amostra por constarem no primeiro semestre do Ensino Superior. O tempo médio de duração da pesquisa foi de 25 minutos.

Resultados

Inicialmente foram realizadas as análises exploratórias, de forma a verificar a adequação dos dados aos pressupostos da análise (Hair Jr, Anderson, Tatham & Black, 2005). Em relação aos casos omissos, foram observados 0,36% de *missing*, uma vez que somente foram aceitos no banco de dados os protocolos com no máximo duas respostas em branco por participante, as quais foram substituídas pelos valores médios nas pontuações individuais. No que diz respeito ao tamanho da amostra, ela também se apresentou adequada ao se presumir que no geral, são necessários minimamente 10 sujeitos por parâmetro estimado.

Não foram encontrados problemas quanto aos pressupostos de multicolinearidade (Hair Jr, Anderson, Tatham, & Black, 2005; Khattree & Naik, 1996) e em relação à normalidade, os índices de assimetria *Skewness* foram positivos e variaram de 0,32 a 6,32, enquanto os valores de achatamento *Kurtosis* variaram entre -1,10 e 70,8, havendo índices positivos e negativos, o que indica um padrão assimétrico aos itens do instrumento. Por essa razão, o procedimento de Análise Fatorial Confirmatória (AFC) mais indicado foi o de máxima verossimilhança (Arbuckle, 2007), o qual foi executado por meio do programa estatístico *Analysis of Moments Structures* (AMOS, 18ªEd), tendo sido examinados, dois modelos fatoriais para a EBADEP-A, uni e bifatorial.

Para a realização dos procedimentos confirmatórios do modelo bifatorial da escala, os itens foram agrupados baseando-se nas sugestões obtidas pelo estudo de Souza (2010), no qual a escala seria dividida em dois fatores, o primeiro agrupando 37 descritores das categorias de humor, cognitivos, sociais, motores, de irritabilidade e ansiedade, e o segundo composto pelos oito itens predominantemente vegetativos ou somáticos. Obviamente, o modelo unifatorial baseado em Baptista (2012) considerou que todos os 45 itens da EBADEP-A são partes de um único fator. Os índices de ajuste dos modelos propostos podem ser observados na Tabela 1.

Inserir Tabela 1 aproximadamente aqui

Foram verificados para o teste de ajuste do modelo (Tabela 1) os índices de Chi-quadrado (X^2), grau de liberdade (gl), bem como a razão X^2/gl , *Goodness-of-fit Index* (GFI), *Normed Fit Index* (NFI), *Comparative Fit Index* (CFI) e *Root Mean Square Error of Approximation* (RMSEA). O modelo bifatorial para a EBADEP-A passou por 15 interações, enquanto o unifatorial passou por 10 interações, sendo considerados como critérios satisfatórios os consensualmente estabelecidos por diferentes autores (Arbuckle, 2007; Cole, 1987; Kline, 2004; Marsh, Balla, & McDonald, 1988; Primi, Muniz & Nunes, 2009). Ambos os modelos apresentaram índice *Root Mean Square Residual* (RMR) igual a 0,04, indicando que não há elementos a serem aprimorados.

Ao se analisar dois modelos à procura da melhor representação do traço latente por um instrumento, todos os índices de ajuste supracitados são comparados. Entretanto um dos ajustes mais importantes envolve a razão X^2/gl (Kline, 2004; Marsh, Balla, & McDonald, 1988). Geralmente, valores entre 2 e 3 para esta razão são recomendados como ajustados,

embora valores próximos de 5 também possam ser considerados como indicadores de ajuste (Kline, 2004). À primeira vista, poder-se-ia julgar o modelo bifatorial como o mais adequado para a EBADEP-A, devido ao valor resultante da razão X^2/df ser levemente o menor. Contudo, percebe-se que ambos estão próximos ao índice máximo de ajuste, o que sugere que não somente os valores brutos devem ser considerados para a escolha do melhor modelo, mas também outras investigações como por exemplo a análise sobre o conteúdo dos fatores.

De modo a explorar quais outros aspectos estariam envolvidos na escolha do modelo mais ajustado para a EBADEP-A, foi realizada uma Análise de Componentes Principais a fim de se observar o comportamento dos dados. Inicialmente, foram gerados 8 fatores, explicando 57,64% da variância total. Contudo, o primeiro fator explicava sozinho 34,9% da variância total, enquanto os demais explicavam entre 2,27% (fator 8) e 4,52% (fator 2), sugerindo estrutura unifatorial. Além disso, a análise do gráfico *Scree Plot* (Figura 1) também apresentou uma estrutura sugerindo um único fator. Quando forçada a estrutura bifatorial para o instrumento, os itens agrupados para cada fator não se organizaram semanticamente válidos, e os descritores somáticos não se agruparam.

Inserir Figura 1 aproximadamente aqui

Assim, tendo em vista que não somente os valores brutos gerados pelas estimativas devem ser considerados na escolha de uma estrutura fatorial, mas também sua organização semântica (Cole, 1987; Kline, 2004), pode-se inferir que a estrutura mais adequada e parcimoniosa do instrumento para a mensuração do traço latente discutido foi o modelo

unifatorial. A Figura 2 apresenta sua reprodução gráfica, a qual foi assumida por Baptista (2012) como a melhor organização fatorial da EBADEP-A.

Inserir Figura 2 aproximadamente aqui

Na análise gráfica da estrutura fatorial da EBADEP-A (Figura 2), cada retângulo no diagrama representa um dos indicadores de depressão presentes na escala, enquanto o construto medido (depressão) é representado pela elipse central. Assim, pode-se observar que embora cada descritor, somático, cognitivo, de humor, ou vegetativo, por exemplo, possuam particularidades que os diferenciam entre si, quando destinados à avaliação do construto em questão, a depressão, eles formam um único grupo.

Discussão

A Escala Baptista de Depressão – Versão Adulto (EBADEP-A) é um instrumento autoaplicável, construída no contexto brasileiro, com intuito de avaliar a sintomatologia depressiva em amostras clínicas e não clínicas (Baptista, 2012). Foram utilizados para a composição dos itens, descritores da depressão contidos tanto nos manuais diagnósticos internacionais quanto nas teorias Cognitivo e Comportamental (APA, 2002; Beck, & cols., 1997; Ferster, Culbertson, & Boren, 1977; OMS, 1993) o que possibilitou a inclusão de um amplo número de descritores de medida.

Embora diferentes instrumentos possam ser encontrados para a mensuração da depressão (Calil & Pires, 1998; Santor, Gregus, & Welch, 2006), ao se comparar os tipos de sintomas presentes na EBADEP-A com os demais, pode-se encontrar uma distribuição diferenciada das categorias sintomáticas avaliadas, quais sejam, de humor, cognitivos,

sociais, vegetativos ou somáticos, motores, de irritabilidade e ansiedade (Baptista, 2012; Calil & Pires, 1998).

Ao longo de sua construção foram estudadas três versões, com números distintos de itens, sendo que desde a primeira delas, com 75 itens, foram publicados diversos estudos com investigações acerca de suas qualidades psicométricas, mais especificamente, diferentes evidências de validade e precisão (Baptista, 2012; Baptista, Souza, & Alves, 2008; Gomes & Baptista, 2010; Souza, 2010; dentre outros). A versão final do instrumento, publicada e autorizada para uso clínico profissional pelo Sistema de Avaliação de Testes Psicológicos (CFP, 2011), com 45 itens também apresenta diferentes estudos verificando evidências de validade de construto, critério, além de sua estabilidade temporal e precisão (Baptista, 2012; Baptista, Cardoso, & Gomes, 2012; Baptista & Gomes, 2011).

Mesmo diante de um número razoável de pesquisas, somente foram encontradas duas as quais buscaram avaliar a organização fatorial do instrumento (Baptista, 2012; Souza, 2010) e por este motivo, o presente estudo teve como objetivo realizar, por meio da análise fatorial baseada no modelo de equações estruturais, a confirmação de qual estrutura da EBADEP-A é a mais adequada ou consistente. Inicialmente, foram atentados os procedimentos de preparação do banco de dados e averiguados os pressupostos para a realização da análise e, em seguida, os modelos de ajuste foram verificados e comparados.

Puderam ser classificados como satisfatórios os índices de ajuste comparativos (CFI), que devem ser entre zero e um; dos índices de bondade de ajuste (GFI), os quais se apresentaram dentro do intervalo requisitado (menores ou iguais a 1); e as raízes da média dos quadrados dos erros de aproximação (RMSEA), valor o qual se almejavam valores menores ou iguais a 0,08. Além disso, foi observado baixo índice de resíduos, medido pelo

RMR, embora os índices de ajuste normalizados (NFI) não tenham ultrapassado 0,08, enquanto esperavam-se valores maiores que 0,90 (Arbuckle, 2007; Kline, 2004).

Especificamente quanto às medidas de ajuste supracitadas, alguns importantes pontos devem ser discutidos e ressaltados. Primeiramente quanto ao CFI, pois embora ele possa ser interpretado como altamente ajustado quando acima de 0,90 as estimativas muito próximas a 1,0 não indicam necessariamente ajuste perfeito dos dados, o que permite a inferência de na presente pesquisa, os valores de 0,72 para o modelo unifatorial e 0,77 para o bifatorial podem ser considerados adequados. Por sua vez, para o GFI, relativo à proporção da covariância explicada pelo modelo, o valor máximo possível, ou seja, 1,0, permite a inferência de ajuste perfeito. Entretanto, deve-se atentar para o fato de que tal índice sofre influência do tamanho da amostra, e por este motivo, embora à primeira vista o valor da organização bifatorial tenha sido levemente maior, ele pode ter recebido influência pelo número de participantes no banco de dados, o que levanta a necessidade de análise de outros índices além destes (Arbuckle, 2007; Cole, 1987; Kline, 2004; Thompson, 2004).

Outro valor amplamente verificado na Análise Fatorial Confirmatória é o RMSEA, uma das formas mais adequadas de se analisar o quanto um modelo se ajusta à população estudada, pois parte do princípio de que não se estuda uma cópia do que se observa, mas sim uma aproximação. Por estar altamente relacionado ao princípio de parcimônia, quando dois modelos são comparados, o mais simples é o favorecido. As estimativas resultantes abaixo de 0,05 são interpretadas como alta aproximação (entre o real e o observado), enquanto valores entre 0,05 e 0,08 são considerados toleráveis, embora possam indicar erros de aproximação (Kline, 2004; Thompson, 2004). Dessa forma, pode-se perceber que, de maneira geral os valores encontrados nas organizações uni e bifatorial foram próximos e

considerados ajustados, embora não possam sozinhos gerar inferências sobre o modelo com maior acurácia.

Foram também comparadas as raízes quadradas médias residuais (RMR), referentes à comparação entre correlações esperadas pelo modelo e as atingidas com base no banco de dados e o NFI, relativo à comparação dos resultados dos chi-quadrados e no qual são desejados valores acima de 0,95. No presente estudo, os valores residuais foram baixos, uma vez que quanto mais próximo de zero, menor a possibilidade de erro, embora os valores de NFI não tenham ultrapassado 0,80 (Arbuckle, 2007; Kline, 2004; Thompson, 2004).

O valor do Chi-quadrado (X^2) é crucial na modelagem de equações estruturais, entretanto ele não pode analisado sem a contextualização anterior dos demais valores, pois assim como o GFI, ele sofre influência do tamanho da amostra (Arbuckle, 2007; Kline, 2004; Primi, Muniz & Nunes, 2009). De acordo com Marsh, Balla e McDonald (1988), ao se considerar somente o valor do chi-quadrado em uma comparação de modelos, pode-se fazer com que a AFC seja contraproducente, gerando interpretações pouco viáveis. Por este motivo, um dos índices principais trata-se da razão X^2/gl uma vez que a divisão do valor bruto pelo grau de liberdade permite a sua normalização em função dos dados (Kline, 2004; Marsh, Balla, & McDonald, 1988). De acordo com Kline (2004) valores entre três e cinco, ou até mesmo em alguns casos, maiores que cinco, podem ser recomendados como indicadores de ajuste.

Os índices encontrados para a razão X^2/gl foram muito parecidos nos dois modelos testados, e muito próximos ao valor máximo de adequação (Kline, 2004), o que pode sugerir a necessidade de reconhecimento de outras análises para se concluir pela

organização fatorial mais adequada. De acordo com Cole (1987), a presença de parâmetros altamente próximos sugere que os componentes constituintes do modelo em que os descritores estão separados em mais de um tipo de medida, isto é, em dois fatores, possuem correlação de tal forma que devam ser apresentados agrupados como um único fator, por ser o modelo mais parcimonioso. O autor considera ainda que tal decisão estatística somente deve ser apurada após uma sistemática revisão das questões teóricas de ambos os modelos. Assim, justificou-se a realização de uma análise exploratória, a qual, pelos critérios de análise de *Scree Plot* e porcentagem de variância explicada (Hair Jr, Anderson, Tatham & Black, 2005) foi confirmada a organização unifatorial como a mais ajustada.

De acordo com Brown (2006), a análise exploratória é limitada a somente a especificar o número de fatores possíveis para uma escala, enquanto na AFC outras soluções são estudadas, a fim de se cumprir, de modo detalhado, o princípio de Parcimônia. Segundo o autor, enquanto a análise exploratória soluciona a presença ou não de fatores, a AFC pode dar suporte à organização fatorial de uma escala, à divisão em diferentes escores ou ao uso do escore total, ou seja, a uma estrutura comporta por todos os itens. Contudo, a análise confirmatória é altamente dependente da amostra a qual se escolhe para a realização dos procedimentos estatísticos, e por esta razão, os parâmetros somente podem ser analisados como satisfatórios ou não e serem generalizados quando os dados assim permitem (Cole, 1987; Hair Jr, Anderson, Tatham & Black, 2005) e na presente pesquisa os valores de simetria e achatamento indicaram um padrão assimétrico dos itens.

Deve-se ressaltar também que não somente as soluções operacionais estatísticas podem ser consideradas durante o processo de construção ou validação de um instrumento. Tanto a análise dos parâmetros envolvidos na AFC quanto as inferências geradas com base

em dados puros, sem a inclusão de embasamentos teórico-críticos, podem gerar um afastamento dos propósitos teóricos e conseqüentemente gerar inferências infundadas ou enganosas (Cole, 1987; Kline, 2004). Como ressalta Pasquali (1998), também podem ser encontradas organizações com boas qualidades estatísticas, mas que não esgotam as características semânticas do construto, o que de fato ocorreu ao se organizar, pelo método de análise de componentes principais, o instrumento em dois fatores.

Dessa forma, o presente estudo apresentou o modelo unifatorial da EBADEP-A como o de maior consistência. Contudo, ressalta-se que embora o unifatorial tenha sido o mais parcimonioso, o modelo bifatorial não deve ser desconsiderado, pois um modelo descartado em um tipo de amostra pode ser investigado em outro conjunto de participantes (Cole, 1987). Estudos futuros devem ser realizados a fim de que as limitações da presente pesquisa possam ser suprimidas. Entretanto, ressalta-se que um novo estudo deve atentar ao tipo e tamanho da amostra, pois podem criar modelos não condizentes com o real medido e/ou gerar modelos tendentes à rejeição (Marsh, Balla, & McDonald, 1988). À guisa de conclusão, pode-se afirmar que as estimativas encontradas deram suporte à interpretação de que a composição unifatorial proposta é coerente com a base teórica do construto avaliado, podendo indicar evidência de validade com base na estrutura interna para a EBADEP-A (Brown, 2006; Hair Jr, Anderson, Tatham, & Black, 2005; Pilati & Laros, 2007; Primi, Muniz, & Nunes, 2009).

Referências

- Abreu, P. R. (2006). Terapia analítico-comportamental da depressão: uma antiga ou nova ciência aplicada? *Revista de Psiquiatria Clínica*, 33(6), 322-328.
- American Psychiatric Association [APA]. (2002). *DSM-IV-TR - Manual diagnóstico e estatístico de transtornos mentais*. (4ª ed). Porto Alegre: Artmed.
- Arbuckle, J. L. (2007). *Amos 18 User's Guide*. Chicago: SPSS Inc.
- Au, A. C. Y., Lau, S., & Lee, M. T. Y. (2009). Suicide ideation and depression: the moderation effects of family coherion and social self-concept. *Adolescence*, 44(176), 851-868.
- Baptista, M. N. (2012). *Manual Técnico da Escala Baptista de Depressão – Versão Adulto (EBADEP-A)*. São Paulo: Vetor.
- Baptista, M. N. (2004). *Suicídio e Depressão: Atualizações*. São Paulo: Guanabara Koogan.
- Baptista, M. N., & Gomes, J. O. (2011). Escala Baptista de Depressão (Versão Adulto) – EBADEP-A: evidências de validade de construto e de critério. *Psico-USF*, 16(2), 151-161.
- Baptista, M. N., Argimon, I. I. L., & Yoshida, E. M. P. (2012). Escalas de Avaliação de Sintomas Depressivos. Em *Avaliação Psicológica nos Contextos Educativo e Psicossocial*. E. Boruchovitch, A. A. A. Santos, & E. Nascimento. São Paulo: Casa do Psicólogo.
- Baptista, M. N., Cardoso, H. F., & Gomes, J. O. (2012). Escala Baptista de Depressão (Versão Adulto) - EBADEP-A: validade convergente e estabilidade temporal. *Psico-USF*, 17(3).

- Baptista, M. N., Souza, M. S., & Alves, G. A. (2008). Evidências de validade entre a Escala de Depressão (EDEP), o BDI e o Inventário de Percepção de Suporte Familiar (IPSF). *Psico-USF, 13*(2), 211-220.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1997). *Terapia da Depressão*. Rio de Janeiro: Zahar.
- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.
- Calil, H. M., & Pires, M. L. N. (1998). Aspectos gerais das escalas de avaliação de depressão. *Revista de Psiquiatria Clínica, 25*(5), 240-244.
- Cavalcante, S. N. (1997). Notas sobre o fenômeno depressão a partir de uma perspectiva analítico-comportamental. *Psicologia, Ciência e Profissão, 17*(2), 2-12.
- Chachamovich, E., Stefanello, S., Botega, N., & Turecki, G. (2009). Quais são os recentes achados clínicos sobre a associação entre depressão e suicídio? *Revista Brasileira de Psiquiatria, 31*(supl. I), 618-625.
- Cole, D. A. (1987). Utility of Confirmatory Factor Analysis in Test Validation Research. *Journal of Consulting and Clinical Psychology, 55*(4), 584-594.
- Conselho Federal de Psicologia [CFP]. (2011). Sistema de Avaliação de Testes Psicológicos (SATEPSI). Disponível em: <http://www.pol.org.br/satepsi/>
- Coutinho, M. P. L., Carolino, Z. C. G., & Medeiros, E. D. (2008). Inventário de Depressão Infantil (CDI): evidências de validade de constructo e consistência interna. *Avaliação Psicológica, 7*(3), 291-300.
- Del Porto, J. A. (1999). Conceito e Diagnóstico. *Revista Brasileira de Psiquiatria, 21* (1), 6-11.

- Del-Ben, C. M., Vilela, J. A. A., Crippa, J. A. S., Hallak, J. E. C., Labate, C. M., & Zuardi, A. (2001). Confiabilidade da “Entrevista Clínica Estruturada para o DSM-IV – Versão Clínica” traduzida para o português. *Revista Brasileira Psiquiatria*, 23, 156-159.
- Ferster, C. B. (1973). A functional analysis of depression. *American Psychologist*, 28, 857-870
- Fleck, M. P. A., Lafer, B., Sougey, E. B., Del Porto, J. A., Brasil, M. A., & Jurena, M. F. (2003). Diretrizes da Associação Médica Brasileira para o tratamento da depressão (versão integral). *Revista Brasileira de Psiquiatria*, 25(2), 114-122.
- Gorenstein, C., & Andrade, L. H. S. G. (1998). Inventário de Depressão de Beck: propriedades psicométricas da versão em português. *Revista de Psiquiatria Clínica*, 25(5), 245-250.
- Grassi-Oliveira, R., Brieztke, E., Teixeira, A., Pezzi, J. C., Zanini, M., Lopes, R. P., & Bauer, M. E. (2012). Peripheral chemokine in women with recurrent major depression with suicidal ideation. *Revista Brasileira de Psiquiatria*, 34(1), 71-75.
- Gomes, J. O., & Baptista, M. N. (2010). Escala de Depressão (EDEP) e medidas de atenção dividida e sustentada em universitários. *Boletim de Psicologia*, 60(133), 191-204.
- Hair Jr., J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise Multivariada de Dados*. Porto Alegre: Bookman.
- Hambleton, R. K., & Jones, R. W. (1993). Comparison of Classical test Theory and Item Response Theory and Their Applications to Test Development. *Educational Measurement: Issues and Practice*, 12(3), 253-262.
- Khattree, R., & Naik, D. N. (1999). Applied Multivariate Statistics with SAS[®] Software (2nd Ed). Cary, NC: SAS Institute Inc.

- Kline, R. B. (2004). *Principles and Practice of Structural Equation Modeling* (2nd Ed). New York: The Guilford Press.
- Matos, E. G., Matos, T. M. G., & Matos, G. M. G. (2006). Depressão Melancólica e Depressão Atípica: aspectos clínicos e psicodinâmicos. *Estudos de Psicologia*, 23(2), 173-179.
- Millis, J. F., & Kroner, D. G. (2008). Predicting Suicidal Ideation with the Depression Hopelessness and Suicide Screening Form (DHS). *Journal of Offender Rehabilitation*, 47(1/2),74-100.
- Moreno, R. A., & Moreno, D. H. (1998). Escalas de depressão de Montgomery & Åsberg (MADRS) e de Hamilton (HAM-D). *Revista de Psiquiatria Clínica*, 25(5), 1-17.
- Organização Mundial da Saúde (OMS) (1993). *Classificação dos Transtornos Mentais e do comportamento - CID-10: descrições e diretrizes diagnósticas*. Trad. Dorgival Caetano. (3º Volume, 10ª Ed.). Porto Alegre: Artes Médicas.
- Panzarella, C., Alloy, L. B., & Whitehouse, W. G. (2006). Expanded Hopelessness Theory of Depression: On the mechanisms by which social support protects against depression. *Cognitive Therapy Ressearch*, 30, 307-333.
- Paradela, E. M. P., Lourenço, R. A., & Veras, R. P. (2005). Validação da escala de depressão geriátrica em um ambulatório geral. *Revista de Saúde Pública*, 39(6), 918-923.
- Pasquali, L. (1998). Princípios de elaboração de escalas psicológicas. *Revista de Psiquiatria Clínica*, 25(5), 206-213.
- Pilati, R., & Laros, J. A. (2007). Modelos de Equações Estruturais em Psicologia: conceitos e aplicações. *Psicologia: Teoria e Pesquisa*, 23(2), 205-216.

- Primi, R., Muniz, M., & Nunes, C. H. S. S. (2009). Definições contemporâneas de Validade de Testes Psicológicos. Em C. S. Hutz (Org). *Avanços e polêmicas em avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Santor, D. A., Gregus, M., & Welch, A. (2006). Eight Decades of Measurement in Depression. *Measurement*, 4(3), 135-155.
- Santos, I. S., Matijasevich, A., Tavares, B. F., Barros, A. J. D., Botelho, I. P., Lapolli, C., Magalhães, P. V. S., Barbosa, A. P. P. N., & Barros, F. C.. (2007). Validation of the Edinburgh Postnatal Depression Scale (EPDS) in a sample of mothers from the 2004 Pelotas Birth Cohort Study. *Cadernos de Saúde Pública*, 23(11), 2577-2588.
- Souza, M. S. (2010). *Evidências de Validade e Precisão para a Escala de Depressão de Baptista e Sisto (EDEP)*. Tese de doutorado. Programa de Pós-Graduação *Stricto Sensu* em Psicologia, Universidade São Francisco, Itatiba, SP.
- Suija, K., Kalda, R., & Maaros, H-I. (2009). Patients with depressive disorder, their co-morbidity, visiting rate and disability in relation to self-evaluation on physical and mental health: a cross-sectional study in family practice. *BMC Family Practice*, 10 (38).
- Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-Fit Indexes in Confirmatory Factor Analysis: The Effect of Sample Size. *Psychological Bulletin*, 103(3), 391-410.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: understanding concepts and applications*. Washington: American Psychological Association.

Tabela 1

Índices de ajuste dos dois modelos testados para a EBADEP-A

| | X^2 | gl | X^2/gl | NFI | GFI | CFI | RMSEA (IC90%) |
|--------------------|--------|-----|-----------------|------|------|------|-----------------------|
| Modelo bifatorial | 4119,7 | 944 | 4,36 | 0,72 | 0,74 | 0,77 | 0,07 (0,071-0,075) |
| Modelo unifatorial | 4435,7 | 945 | 4,69 | 0,70 | 0,72 | 0,72 | 0,08 (0,074-0,079) |

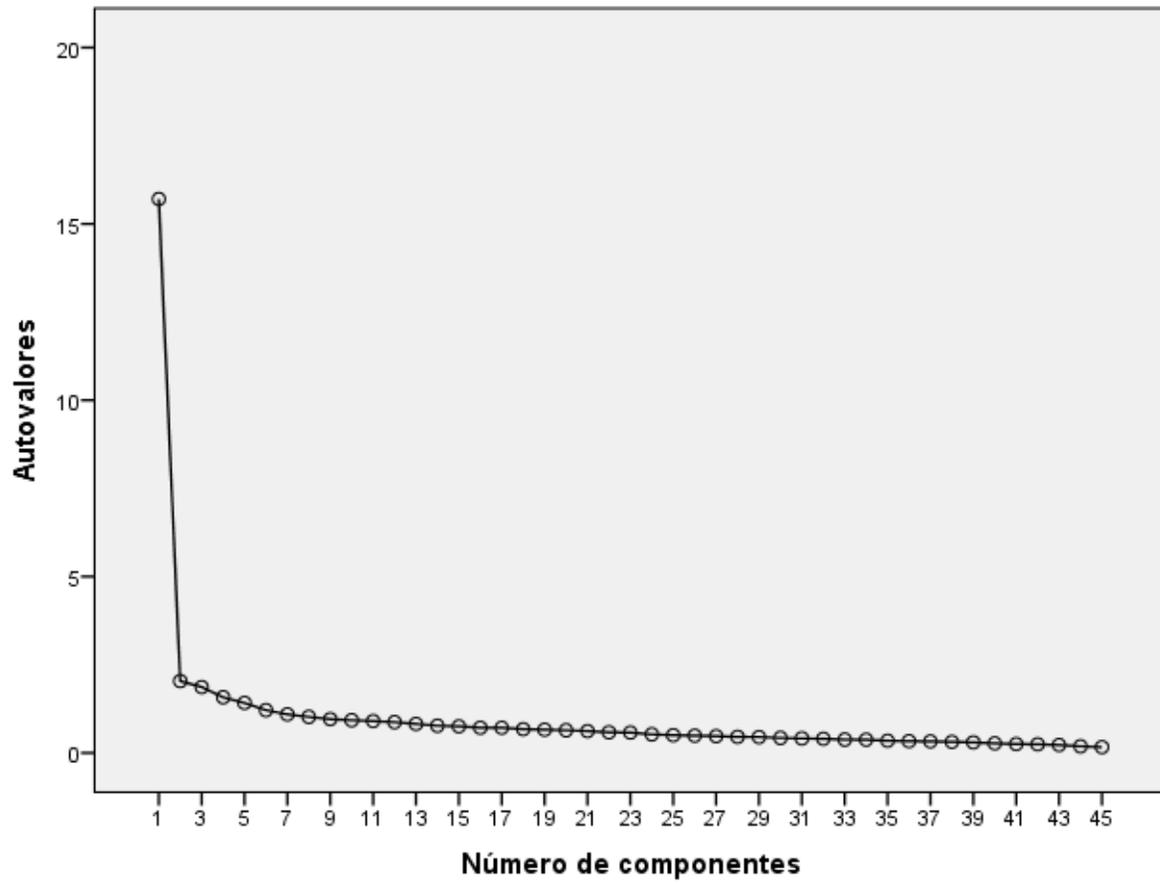


Figura 1. Gráfico *Scree Plot*

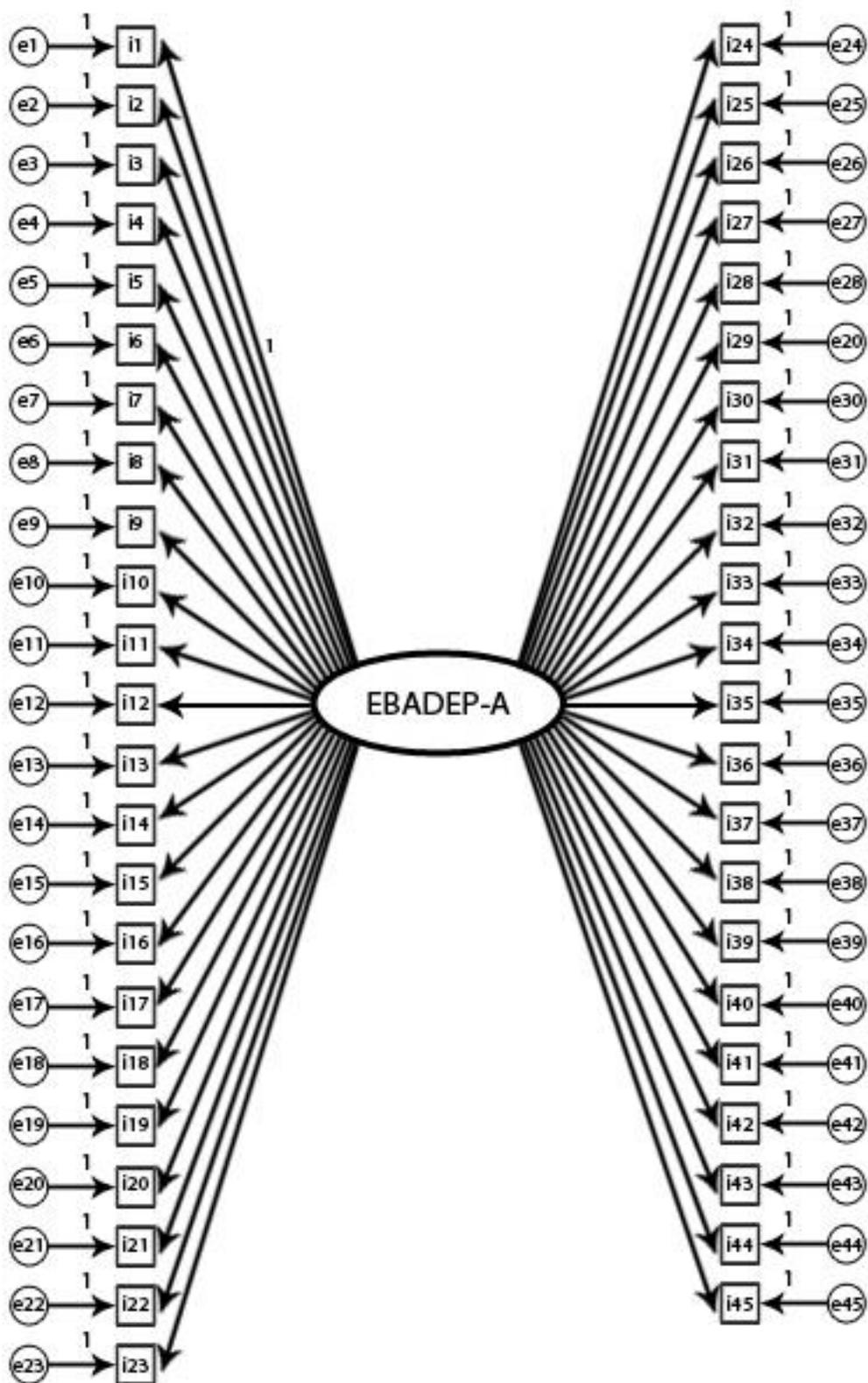


Figura 2. Modelo unifatorial da EBADEP-A

Artigo 3

Normalization procedure for the Baptista Depression Scale - Adult Version

(EBADEP-A): transferring of norms

Regarding the standardization of psychological assessment instruments, that is, the construction of referential interpretation of a test, different procedures can be found performed by Classical Test Theory and the Theory of Item Response. Especially in this case, it can be sustained a standardization of a test as a pattern to transfer its cut-off point to another instrument. Thus, the present study aimed to provide a cutoff score for the Baptista Depression Scale - Adult Version (EBADEP-A) through procedures of norms-transfer based on the Center for Epidemiologic Studies – Depression Scale (CES-D). The EBADEP-A presented good distribution and ability to discriminate depressive symptoms, and received a cutoff score of 32 point to the sample, consisting of Brazilian College students. It is emphasized that this is an exploratory and preliminary study, and it is suggested further analyzes to be performed with clinical samples for which results can be corroborated or confronted.

Keywords: psychological assessment; test interpretation; psychometrics

Procedimento de normatização para a Escala Baptista de Depressão – Versão Adulto (EBADEP-A): transferência de normas

No que concerne à normatização de instrumentos de avaliação psicológica, ou seja, a construção dos referenciais de interpretação de um teste, diferentes procedimentos podem ser encontrados, realizados com base na Teoria Clássica dos Testes e Teoria de Resposta ao Item. Especialmente neste caso, pode-se ostentar uma normatização de um teste como padrão, para transferir seus pontos de corte a outro instrumento. Assim, o presente estudo teve como objetivo apresentar uma pontuação de corte para a Escala Baptista de Depressão – Versão Adulto (EBADEP-A), por meio de procedimentos de transferência de normas com base na *Center for Epidemiologic Studies – Depression Scale* (CES-D). A EBADEP-A apresentou boa distribuição e capacidade de discriminar a sintomatologia depressiva, e recebeu a pontuação de corte igual a 32 para amostra universitária. Ressalta-se que se trata de um estudo exploratório e preliminar, e sugere-se que novas análises com amostras clínicas sejam realizadas para que os resultados sejam corroborados ou confrontados.

Palavras-chave: avaliação psicológica; interpretação do teste; psicometria

Procedimiento de normalización de la Baptista Escala de Depresión - Versión Adulto

(EBADEP-A): transferencia entre estándares

En cuanto a la normalización de los instrumentos de evaluación psicológica, es decir, la construcción de la interpretación referencial de una prueba, distintos procedimientos se pueden encontrar, hecho basado en la Teoría Clásica de los Tests y la Teoría de Respuesta al Ítem. Especialmente en este caso, puede ser sostenida una prueba como estándar y realizar una transferencia de límites de puntuación a otro instrumento. Así, el presente estudio tuvo como objetivo proporcionar un punto de corte para la Baptista Escala de Depresión - Versión Adulto (EBADEP-A), por procedimientos de transferencia de estándares basado en la Center for Epidemiologic Studies – Depression Scale (CES-D). La EBADEP-A mostró buena distribución y capacidad de discriminar los síntomas depresivos, y recibió un puntaje de corte igual a 32 para la muestra de universitarios. Se enfatiza que se trata de un estudio exploratorio y preliminar, y se sugiere que adicionalmente se realizan análisis con muestras clínicas para que los resultados se corroboren o sean confrontados.

Palabras clave: evaluación psicológica; interpretación de test; psicometría

When **regarding to** the construction of psychological instruments, it is generally discussed the basic psychometric characteristics required for them to be considered suitable for use. In Brazil, the use of psychological tests is restricted to the Psychologist, and so to ensure the minimal quality of the tests, questionnaires, scales or inventories used, the Brazilian Psychological Council, by means of a resolution (Psychologist Federal Council [CFP], 2003), gathered the basic psychometric criteria for psychological instruments based on the International Test Commission (ITC, 2000) and on the Standards for Educational and Psychological Test (American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME], 1999). Specifically, there are two psychometric qualities that can be investigated, known as reliability and validity (Anastasi & Urbina, 2000; Noronha & Alchieri, 2004; Noronha & Vendramini, 2003; Pasquali, 1999; Urbina, 2004).

Reliability is committed to measurement error of the instrument and its stability, **that is,** how much the measurement approaches to or moves away from the factual characteristics of the individual. So, the higher the accuracy of a test, greater its reliability, and lower the error in the measurement. In turn, validity **regards** the assumption that the test, **may** its questions, has the ability to relate the construct which it is involved (Anastasi & Urbina, 2000; Pasquali, 1998; 1999; Urbina, 2004).

The use of a reliable and valid instrument enables the researcher, as well as the psychologist, to survey different psychological issues; even those not-so-easily observed (Noronha, 2009). Within this perspective of evaluation, named Classical Test Theory (CTT), the characteristics of the sample is **involved** in reliability and validity studies, are influential factors to the statistical results, and because of that it is possible to find different

options of reliability indices or distinct validity evidences, according to each sample investigated.

Meanwhile, considering both the historical criticism that the Psychological Assessment received over the years in Brazil, with its importance been disregarded and reconsidered for several reasons, as the ever-present need of theoretical and methodological updates, different alternative methods for the CTT procedures can be found (Amarnani, 2009; Hambleton & Jones, 1993; Pasquali & Primi, 2003). One of the theories, which suggested the study and interpretations of a psychological instrument to be changerground, was the Item Response Theory (IRT) (Hambleton & Jones, 1993; Nunes & Primi, 2009; Pasquali & Primi, 2003; Valentini & Laros, 2011).

On the one hand, the CTT provides a methodological focus on the characteristics of the sample, since for each group investigated distinct reliability indices can be verified, and different evidences of validity, noting that reliability and validation do not decay on the test itself, but to the result interpretations of the Theory *versus* Results relationship (Urbina, 2004). On the other hand, the IRT presents a set of representation models of the test parameters, focusing both on the characteristics of the items, and in the probability of a person to choose one or another answer according to one's skill level, emphasizing the latent trait (Amarnani, 2009; Embretson & Reise, 2000; Pasquali & Primi, 2003; Rueda, 2007; Valentini & Laros, 2011).

Naturally, several comparisons between CTT and IRT are expected to be found in the literature, showing their operation differences, advantages and disadvantages (Amarnani, 2009; Embretson & Reise, 2000; Fan, 1998; Hambleton & Jones, 1993; Pasquali & Primi, 2003; Rueda, 2007; Valentini & Laros, 2011; Wiberg, 2004; among

others). As a result of these various comparisons, it can be said that although the development of tests based on CTT cannot be considered a negative characteristic for Psychological Assessment, tests can indeed be benefited with the use of most modern and advanced statistical methods and models, as the IRT (Baptista & Gomes, 2011; Forkmann et al., 2009). Another trait that may be cited refers to the programs available for performing the calculations, which enable to verify the residue index obtained between what was expected, and the result obtained (Nunes & Primi, 2009).

During the standardization of test procedures, i.e. the construction of the referential interpretation of the instrument (Urbina, 2004), differences related to the CTT and IRT are also found. In the first one, the procedures take into consideration the sample normal distribution, and the results interpretation is connected to a main group of scores, called normative, under which the standards are constructed (Anastasi & Urbina, 2000; Urbina, 2004). In the second one, in turn, assuming the focus on latent trait and on the instrument, we can admit the standardization of a test by considering one instrument as the default, and later transfer its cutting points to the other one (Thomas, 2011).

For such a procedure, items and persons present in the database are subjected to a calibration, so that the parameters are separated, and therefore they lose their dependence on each other. Furthermore, measurements of both instruments are also calibrated on a common scale, allowing the comparison and the transfer of normative standards from one instrument to another, through the equalization procedure (Bauer & Hussong, 2009; Embretson & Reise, 2000; Thomas, 2011). The equalization allows instruments which theoretically measure the same construct, to become comparable and proportionial, making

it possible that the same statistical significance (and so their interpretations), to be attributed to participants with the same ability (Smith et al., 2006; Wyse & Reckase, 2011).

Although IRT is presented with a predominant focus on skills and ability instrument, like trial-and-error instruments, studies which use these models to assess more subjective constructs can be also found, related to mental health, such as depression for example (Castro, Trentini, & Riboldi, 2010; Covic, Pallant, Rabin, & Kaufman, 2004; Cole et al., 2011; Covic, Pallant, Cnaghan, & Tennant, 2007; Jones & Fonda, 2004; Pickard, Dalal, & Bushnell, 2006; Sauer, Ziegler, & Schmitt, 2012; among others). IRT methods can, indeed, provide many benefits to the studies of the most subjective or complex psychopathological measures (Reise & Waller, 2009).

Concerning the depression measurement, several instruments can be presented, many of which with various evidences of validity and reliability in many countries, including Brazil (Calil & Pires; 1998; Santor, Gregus & Welch, 2006). A widely used instrument is the Center for Epidemiologic Studies – Depression Scale (CES-D) which assesses the construct through 20 items, with an emphasis on affective components and depressed mood (Radloff, 1977). Specifically in Brazil, different surveys about validity evidences can be found, from which distinct cutoff points were established for the scale, according to the samples studied, that is, college students (cutoff 15), drugs addicted groups (cutoff 16) and general population (cutoff 24) (Batistoni, Neri & Cupertino, 2007; 2010; Hauck Filho & Teixeira, 2011; Silveira & Jorge, 1998).

The use of imported and translated instruments, since they possess the necessary adaptation, reliability and validity studies, with different reference samples, can benefit both the scientific community and clinicians, who can use them in different contexts

(Urbina, 2004). However, it is always important to underline the importance of not only updating the imported studies with these tests, but also the construction of new instruments, with contextualized items intended for different populations.

In this direction, it is highlighted the Brazilian instrument named Baptista Depression Scale - Adult Version (*Escala Baptista de Depressão – Versão Adulto*; EBADEP-A), which focuses on the assessment of depressive symptoms based on 26 descriptors, over 45 items (Baptista, 2012). The EBADEP-A has evidences of validity and reliability based both in CTT and IRT, and it also presents ratings of the symptoms severity, namely minor/ none symptomatology of depression, mild, moderate and severe symptomatology, based on the transfer of norms procedure from the Beck Depression Inventory (Baptista, 2012; Baptista & Gomes, 2011; Baptista, Cardoso, & Gomes, 2012).

This study assumes that the IRT allows, more accurately than the CTT, the reduction of errors on measurements in the analysis of latent variables. Moreover, it is assumed that it is possible to assess the bias and to analyze the adjustments of the items of these instruments, as well as use calibration and equalization of these measures, so that they might be comparable (Embretson & Reise, 2000; Thomas, 2011). Thus, the present study aimed to provide a cutoff score for the Baptista Depression Scale - Adult Version, through transfer procedures of norms from the Center for Epidemiologic Studies – Depression Scale (CES-D).

Method

Participants

The study included 589 college students from different towns in two Brazilian states, São Paulo and Minas Gerais. From this total, 519 (88%) were from different undergraduate courses, such as Nursing, Geography, Psychology, Law School, Biology, Nutrition, Pharmacy and Physiotherapy, while 43 (7.4%) were from postgraduate courses in Psychology and 27 (4.6%) did not respond. The sample was composed mostly of women ($n = 439$, 74.5%) and ages ranged from 17 to 63 years with mode equal to 18 ($M = 23.48$, $SD = 8.56$). Regarding marital status, the majority were single ($n = 495$, 83.9%).

Instruments

Baptista Depression Scale - Adult Version (EBADEP-A)

The EBADEP-A was constructed by Baptista (2012), from 26 descriptors of depressive symptoms, based on the international diagnostic manuals CID 10 and DSM-IV-TR, as well as in Cognitive and Behavioral theories of depression (APA, 2002; Beck, Rush, Shaw, & Emery, 1997; Ferster, Culbertson & Boren, 1977; OMS, 1993). It has 45 four-point Likert items consist of pairs of sentences with positive and negative imprint which refer to the descriptors of depression. Its minimum score is zero and the maximum, 135, and the higher its score, the greater depression symptoms presented.

Center for Epidemiologic Studies – Depression Scale (CES-D)

It is a screening scale of depressive symptomatology, composed of 20 four-point Likert items. It has been built based on different instruments for depression, in order to be

used both in clinical and non-clinical contexts. Its minimum score is zero and the maximum, 60, and the higher your overall score, the greater presence of depressive symptoms (Radloff, 1977). In Brazil, it was translated by Silveira and Jorge (1998), who presented scores 15 as minimum for the presence of depressive symptoms in college students, 16 for substance abusers and 24 for the general population.

Procedures

To perform this study, all ethical procedures were adopted. The project was approved by the Ethics Committee of the São Francisco University, Brazil, and each participant signed in two copies, the Term of Consent before answering the questionnaires. An Identification Questionnaire was presented to the sample, followed by the EBADEP-A, and the CES-DA. The instruments were answered in approximately 35 minutes. As inclusion criteria, were only accepted protocols which presented minimally 80% of the instruments properly answered. The instruments were answered in approximately 35 minutes.

Results

First, we checked the inclusion criteria and were excluded protocols with number greater or equal to three answers blank on one of the two instruments. As regards to the EBADEP-A, the question with more missing data was on completing tasks ($n = 13$), that is, linked to the descriptors of feelings of inadequacy and loss of productivity, followed by issues related to loss of libido ($n = 8$), helplessness, neediness, depressed mood, lack of perspective with the present ($n = 6$), and fatigue/ energy loss ($n = 4$). In turn, for the CES-D

descriptor with the largest number of blank answers was the lack of perspective with the future ($n = 6$), followed by helplessness ($n = 5$), irritability, and anhedonia ($n = 4$). All questions blank were replaced by individual averages.

Then, in a descriptive way, the structures of each instrument, and adequacy of the data to the model were verified. Regarding to items, infit mean values of the EBADEP-A ($M = 1.03$, $SD = 0.26$) and the CES-D ($M = 1.00$, $SD = 0.32$), and the means of the outfits to the EBADEP-A ($M = 1.02$, $SD = 0.28$) and the CES-D ($M = 1.02$, $SD = 0.51$) indicate that they're being answered in a default line and are suitable to the Rasch model (Linacre, 2010). Specifically about the distribution of the response options, it is presented in the Figure 1 the graphical display related to both scales.

Insert Figure 1 approximately here

It can be seen from Figure 1 that, although the scales have the same Likert options (from zero to three), the distribution does not occur in the same way, as the thresholds between the options are different. To the EBADEP-A, they ranged from -2.52 to 2.32 , while for the CES-D the variation was between -1.83 and 1.83 . Thus, it was revealed that the instruments and the data showed fit to the Rasch model, what demonstrated the possibility of performing the transfer procedures.

In order to proceed, both scales were analyzed together, treated as a single set of items. Through of a map of items (Table 1) it was possible to explore such structure based on the graphical display of skill levels of persons, and with regard to the items difficulty, that is, regarding to the assessment of people which showed more or less severe depression, according to the instruments scores.

Insert Table 1 approximately here

The map of items (Table 1) shows the variation and distribution of the items difficulties, as well as the ability levels of persons. The letters M located in the center column indicate the average positions. It was observed that the average value of the items was larger than the average value of persons, and therefore, the items were rated as more difficult than the abilities. Since EBADE-A and CES-D are not ability instruments, with right-and-wrong questions, it can be said that most questions received the minimum values possible, and all the items were positioned above the average of persons.

First, the item considered more difficult was related to suicidal ideation (EBADEP29), what shows that the participants had chosen most frequently for this item the circles next to the positive sentence in the instrument. Then, three items of the CES-D, referring to perspectives about the past, present and future (CESD16, CESD8, and CESD12), were easier, i.e., have been widely reported with the maximum values. In a general context, it was concluded that although EBADEP-A has more than twice numbers of items than the CES-D, this one presented easier items, with greater possibility of maximum responses.

In order that the instruments could be analyzed on the same scale, it was necessary to calibrate and equalize the measures of test parameters, so that the Theta values could be used to the transfer of CES-D cut-off points to the EBADEP-A (Bauer & Hussong, 2009; Embretson & Reise, 2000; Smith et al., 2006; Thomas, 2011; Wyse & Reckase, 2011). After the procedures of equalization and anchorage, all possible CES-D scores (from zero

to 60) were associated with an ability level from which one could verify the equivalent expected measure to the cutoff point to this sample, of college students (Table 2).

Insert Table 2 approximately here

Thereby, in Table 2, it was possible to find out that the Theta value associated to the cutoff point 15, given by Silveira and Jorge (1998) in their research, is -1.17. Proceeding, in order to transfer such CES-D standard to EBADEP-A, the possible values of this scale (from zero to 135) were also associated with levels of ability (Table 3). In this case, it was searched the Theta measure equivalent to the CES-D cutoff score above mentioned, at the Theta levels shown for EBADEP. For the Theta ability -1.17 it has been found the associated score of 32 points.

Insert Table 3 approximately here

Additionally, through crosstabulation it was possible to indicate that out of 269 participants indicated with lack of depressive symptoms by CES-D, 222 (82.5%) presented the same diagnosis by EBADEP-A, while the coincidence of depression diagnosis occurred in 65% (Table 4). It was also determined the association between variables, using the chi-square test ($X^2 = 134,47$; $df = 1$; $p \leq 0,001$) and Pearson's correlation ($r = 0,71$; $p \leq 0,001$).

Insert Table 4 approximately here

Finally, it is emphasized that the cutoff point of the CES-D for college students (equal to 15) equals to 25% of the total possible score of the instrument. After the transfer of norms, the EBADEP-A she received cutoff score for college students equal to 32, i.e.

23.7% of the possible score. Originally, 320 students (54.3%) had depressive symptoms according to the CESD-D cut-off score. After the transfer of norms, it was found that 255 college students (43.3%) could be classified as depressed by the EBADEP-A.

Discussion

Within the context of psychological instruments construction, it is always emphasized their basic psychometric qualities, namely, the validity and reliability, so that it can be used in different contexts (Anastasi & Urbina, 2000; Noronha & Alchieri, 2004; Noronha & Vendramini, 2003; Pasquali, 1999; Urbina, 2004). For both procedures and techniques, one can use statistical models based in both Classical Test Theory (CTT) and the Item Response Theory (IRT).

The same argument applies to procedures of standardization, i.e. the indication of interpretive rules of the test (Urbina, 2004). On one hand, the CTT is based on the characteristics of sample groups, while on the other hand, the IRT focuses on the characteristics of the items, and the probability of participants in sample to indicate one or another kind of response, with emphasis on the latent trait (Amarnani, 2009; Embretson & Reise, 2000; Pasquali & Primi, 2003; Rueda, 2007; Valentini & Laros, 2011).

Indeed, there are different studies which seek to show the advantages, disadvantages and differences between the CTT and the IRT (Amarnani, 2009; Embretson & Reise, 2000; Fan, 1998; Pasquali & Primi, 2003; Rueda, 2007; Valentini & Laros, 2011; Wiberg, 2004; among others). However, it is noteworthy that although the use of CTT in Psychological Assessment is not a downside to the evolution of the clinical and nonclinical Psychology, and for the progress of the surveys, the use of the techniques shown by models of IRT can

benefit it by using more advanced statistical methods, which present a reduced number of noise and waste (Baptista & Gomes, 2011; Nunes & Primi, 2009; Forkmann et al., 2009).

One of the possible procedures for the standardization of an instrument by IRT involves the transfer of norms of a standards previously valid and reliable test, by comparing the level of ability required to achieve the cutoff score set (Thomas, 2011). For this, instruments must be measured within the same scale, i.e equalized, and subsequently calibrated, in order to make it possible to transfer each instruments into Theta.

Under these assumptions, the aim of this study was to realize a transfer of norms from the Center for Epidemiologic Studies – Depression Scale (CES-D), a rating scale of depressive symptoms widely known and published in different countries, including Brazil (Batistoni et al., 2007; 2010; Hauck Filho & Teixeira, 2011; Radloff, 1977; Silveira & Jorge, 1998), to the Baptista Depression Scale - Adult Version (EBADEP-A), which was built in Brazil and presents different studies of validity and reliability in the country (Baptista, 2012; Baptista & Gomes, 2011; Baptista et al., 2012).

Initially, the model adequation indices were verified by the average values of infits and outfit (Linacre, 2010). It was also observed that both tests showed good distribution of their response options. Although both instruments involved in the study have the same type of Likert scale, between zero and three points, the structures are, indeed, very different, and for this reason, for the transfer of cutoff points, the initial procedures of equalization and anchorage were conducted.

Through the interpretation of the map items, in which both scales were evaluated together, one can see that the items were considered difficult to the sample, since most of the items were located above the population average. This trait indicated that there was

greater adherence of the sample for answering lower choices, near zero, suggesting low depressive symptoms of the persons, which is somewhat expected in a nonclinical population. It can be said that the EBADEP-A is effectively evaluating depressive symptoms in college students, once if the extent of items and people were very similar, the scale would be classified as very easy for the sample, and consequently, it would not be captured adequately depressive symptoms.

A similar procedure of transfers of norms was accomplished by Baptista (2012), based on the Beck Depression Inventory (BDI). Most items of EBADEP-A were below the average of the items and above average of persons, showing that the EBADEP-A presented the ability to differentiate moderate from severe symptoms, and proving that the BDI differed better milder symptoms from the moderate. In the present research, although EBADEP-A nature has not exclusive screening, as with the CES-D, it showed a good ability to track depressive symptoms in college students.

Through the data equalization, both instruments employed in the present study began to exhibit the same scale of difficulty, and therefore, the ability levels (Theta) could be interpreted as equivalent. Initially were set every possible scores of CES-D, in order to find the Theta equivalent to the scale cut-off point. Later, the same procedure was performed for the EBADEP-A, so that from Theta established in the previous procedure, it could be established the cutoff score.

Thus, it can be said that the EBADEP-A showed good screening feature of depressive symptoms in the sample, and it had received a cutoff equal to 32 to the presence of depressive symptoms. However, it should be highlighted that this is an exploratory study, and therefore the results should be treated cautiously. First of all, about the CES-D, it

does not provide points of severity of depressive symptoms, as in the BDI or EBADEP-A, but only one cutoff point, which may limit the interpretations. Furthermore, it raises questions about the sample type used, i.e. participants whose clinical diagnosis of depression was not measured: selecting a sample with a low frequency of symptoms can bring limitations to standardization of the instrument? Future researches are suggested, with a sample of people diagnosed with depressive disorder, in order to corroborate or compare the results discussed here.

References

- Amarnani, R. (2009). Two Theories, One Theta: A Gentle Introduction to Item Response Theory as an Alternative to Classical Test Theory. *The International Journal of Educational and Psychological Assessment*, 3, 104-109.
- American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME]. (1999). *Standards for Psychology and Educational Testing*. Washington, DC: American Psychology Association.
- American Psychiatric Association [APA]. (2002). *DSM-IV-TR - Manual diagnóstico e estatístico de transtornos mentais*. (4^a ed). Porto Alegre: Artmed.
- Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica*. Veronese, M.A.V. (trad.). Porto Alegre: Artmed Editora.
- Baptista, M. N. (2012). *Manual Técnico da Escala Baptista de Depressão – Versão Adulto (EBADEP-A)*. São Paulo: Vetor.
- Baptista, M. N., Cardoso, H. F., & Gomes, J. O. (2012). Escala Baptista de Depressão (Versão Adulto) - EBADEP-A: validade convergente e estabilidade temporal. *Psico-USF*, 17(3).
- Baptista, M. N., & Gomes, J. O. (2011). Escala Baptista de Depressão (Versão Adulto) – EBADEP-A: evidências de validade de construto e de critério. *Psico-USF*, 16(2), 151-161.
- Batistoni, S. S. T., Néri, A. N., & Cupertino, A. P. (2010). Validade e confiabilidade da versão Brasileira da Center for Epidemiological Scale – Depression (CES-D) em idosos Brasileiros. *Psico-USF*, 15(1), 13-22.

- Batistoni, S. S. T., Néri, A. N., & Cupertino, A. P. F. B. (2007). Validade da escala de depressão do Center for Epidemiological Studies entre idosos Brasileiros. *Revista de Saúde Pública, 41*(4), 598-605.
- Bauer, D. J., & Hussong, A. M. (2009). Psychometric Approaches for Developing Commensurate Measures Across Independent Studies: Traditional and New Models. *Psychological Methods, 14* (2), 101-125.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1997). *Terapia da Depressão*. Rio de Janeiro: Zahar.
- Calil, H. M., & Pires, M. L. N. (1998). Aspectos gerais das escalas de avaliação de depressão. *Revista de Psiquiatria Clínica, 25*(5), 240-244.
- Castro, S. M. J., Trentini, C., & Riboldi, J. (2010). Item response theory applied to the Beck Depression Inventory. *Revista Brasileira de Epidemiologia, 13*(3), 1-13.
- Cole, D. A., Martin, N. C., Youngstrom, E. A., Curry, J. F., Essex, M. J., Goodyer, I. M., Stark, K. D., Cai, L., Findling, R. L., Garber, J., Hyde, J. S., Compas, B. E., Rohde, P., Rohde, P., Slaterry, M. J., & Forehand, R. (2011). Structure and Measurement of Depression in Youths: Applying Item Response Theory to Clinical Data. *PsychologicalAssesment, 23*(4), 819-833.
- Cole, J. C., Smith, T. L., Rabin, A. S., Kaufman, A. S. (2004). Development and Validation of a Rasch-Derived CES-D Short Form. *Psychological Assessment, 16*(4), 360-372.
- Conselho Federal de Psicologia – CFP (2003). Resolução nº 002/2003 [On-line]. Retrieved from
http://www.pol.org.br/pol/export/sites/default/pol/legislacao/legislacaoDocumentos/resolucao2003_02.pdf

- Covic, T., Pallant, J. F., Conaghan, P. G., Tennant, A. (2007). A longitudinal evaluation of the Center for Epidemiologic Studies-Depression scale (CES-D) in a Rheumatoid Arthritis Population using Rasch Analysis. *Health and Quality of Life Outcomes*, 6, 6-41.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah: Lawrence Erlbaum.
- Fan, X. (1998). Item response theory and classical test theory: an empirical comparison of their item/person statistics. *Educational and Psychological Measurement*, 58(3), 357-381.
- Ferster, C. B., Culbertson, S., & Boren, C. P. (1977). *Princípios do comportamento*. (trad: Maria Ignez Rocha e Silva, Maria Alice de Campos Rodrigues e Maria Benedita Lima Pardo). São Paulo: Hucitec.
- Forkmann, T., Boecker, M., Wirtz, M., Eberle, N., Westhofen, M., Schauerte, P., et al. (2009). Development and validation of the Rasch-based Depression Screening (DESC) using Rasch analysis and structural equation modelling. *Journal of Behavior Therapy and Experimental Psychiatry*, 40 (3), 468-478.
- Hambleton, R. K., & Jones, R. W. (1993). An NCME Instructional Module on Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Educational Measurement: Issues and Practice*, 12(3), 38-47.
- Hauck Filho, N., & Teixeira, M. A. P. (2011). A estrutura fatorial da escala CES-D em estudantes universitários brasileiros. *Avaliação Psicológica*, 10(1), 91-97.

- International Test Commission – ITC (2000). Guidelines on Adapting Test. *International Test Commission*. [On-Line]. Retrieved from <http://www.intestcom.org/upload/sitefiles/40.pdf>
- Jones, R. N., Fonda, S. J. (2004). Use of an IRT-based latent variable model to link different forms of the CES-D from the Health and Retirement Study. *Social Psychiatry and Psychiatric Epidemiology*, 39, 828-835.
- Linacre, J.M. (2010). Winsteps® (Version 3.70.0.2) [Computer Software]. Beaverton, Oregon: Winsteps.com.
- Noronha, A. P. P. (2009). Testes Psicológicos: conceito, uso e formação do psicólogo. Em Claudio S. Hutz (org). *Avanços e polêmicas em avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Noronha, A. P. P., & Alchieri, J. C. (2004). Conhecimento em Avaliação Psicológica. *Estudos de Psicologia*, 21(1), 43-52.
- Noronha, A. P. P., & Vendramini, C. M. M. (2003). Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica*, 16(1), 177-182.
- Nunes, C. H. S. S., & Primi, R. (2009). Teoria de resposta ao item: conceitos e aplicações na psicologia e na educação. In Hutz, C S (org). *Avanços e polêmicas em avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Organização Mundial da Saúde (OMS) (1993). *Classificação dos Transtornos Mentais e do comportamento - CID-10: descrições e diretrizes diagnósticas*. Trad. Dorgival Caetano. (3º Volume, 10ª Ed.). Porto Alegre: Artes Médicas.
- Pasquali, L. (1998). *Psicometria: Teoria e aplicações*. Brasília: Editora UnB.

- Pasquali, L. (1999). *Instrumentos psicológicos: manual prático de elaboração*. Brasília: LabPAM & IBAPP.
- Pasquali, L., & Primi, R. (2003). Fundamentos da teoria da resposta ao item: TRI. *Avaliação Psicológica*, 2(2), 99-110.
- Pickard, A. S., Dalal, M. R., & Bushnell, D. M. (2006). A Comparison of Depressive Symptoms in Stroke and Primary Care: Applying Rasch Models to Evaluate the Center for Epidemiologic Studies-Depression Scale. *Value in Health*, 1(9), 59-64.
- Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, 1(3), 385-401.
- Rueda, F. J. M. (2007). O funcionamento diferencial do item no teste pictórico de memória. *Revista Avaliação Psicológica*, 6(2), 229-237.
- Santor, D. A., Gregus, M., & Welch, A. (2006). Eight Decades of Measurement in Depression. *Measurement*, 4(3), 135-155.
- Sauer, S., Ziegler, M., & Schmitt, M. (2012). Rasch analysis of a simplified Beck Depression Inventory *Personality and Individual Differences*, in press. doi: 10.1016/j.paid.2012.10.025
- Silveira, D. X., & Jorge, M. R. (1998). Propriedades psicométricas da escala de rastreamento populacional para depressão CES-D em populações clínica e não-clínica de adolescentes e adultos jovens. *Revista de Psiquiatria Clínica*, 25(5), 251-61.
- Smith, A. B., Wright, E. P., Rush, R., Stark, D. P., Velikova, G., & Selby, P. J. (2006). Rasch analysis of the dimensional structure of the Hospital Anxiety and Depression Scale. *Psycho-Oncology*, 15, 817-827.

- Thomas, M. L. (2011). The Value of Item Response Theory in Clinical Assessment: A Review. *Assessment, 18*(3), 291-307.
- Urbina, S. (2004). *Essentials of Psychological Testing*. New Jersey: John Wiley & Sons, Inc.
- Valentini, F., & Laros, J. A. (2011). Teoria de Resposta ao Item na Avaliação Psicologia. In R. A. M. Ambiel, I. S. R., S. V. Pacanaro, G. A. S. Alves, I. F. A. S. Leme (Orgs). *Avaliação Psicológica: guia de consulta para estudantes e profissionais de psicologia*. São Paulo: Casa do Psicólogo.
- Wiberg, M. (2004). Classical Tests Theory vs. Item Response Theory: an evaluation of the theory test in the Swedish driving-license test. *Working Paper EM, 50*, 1-30.
- Wyse, A. E., & Reckase, M. D. (2011). A Graphical Approach to Evaluating Equating Using Test Characteristic Curves. *Applied Psychological Measurement, 35* (3), 217-234.

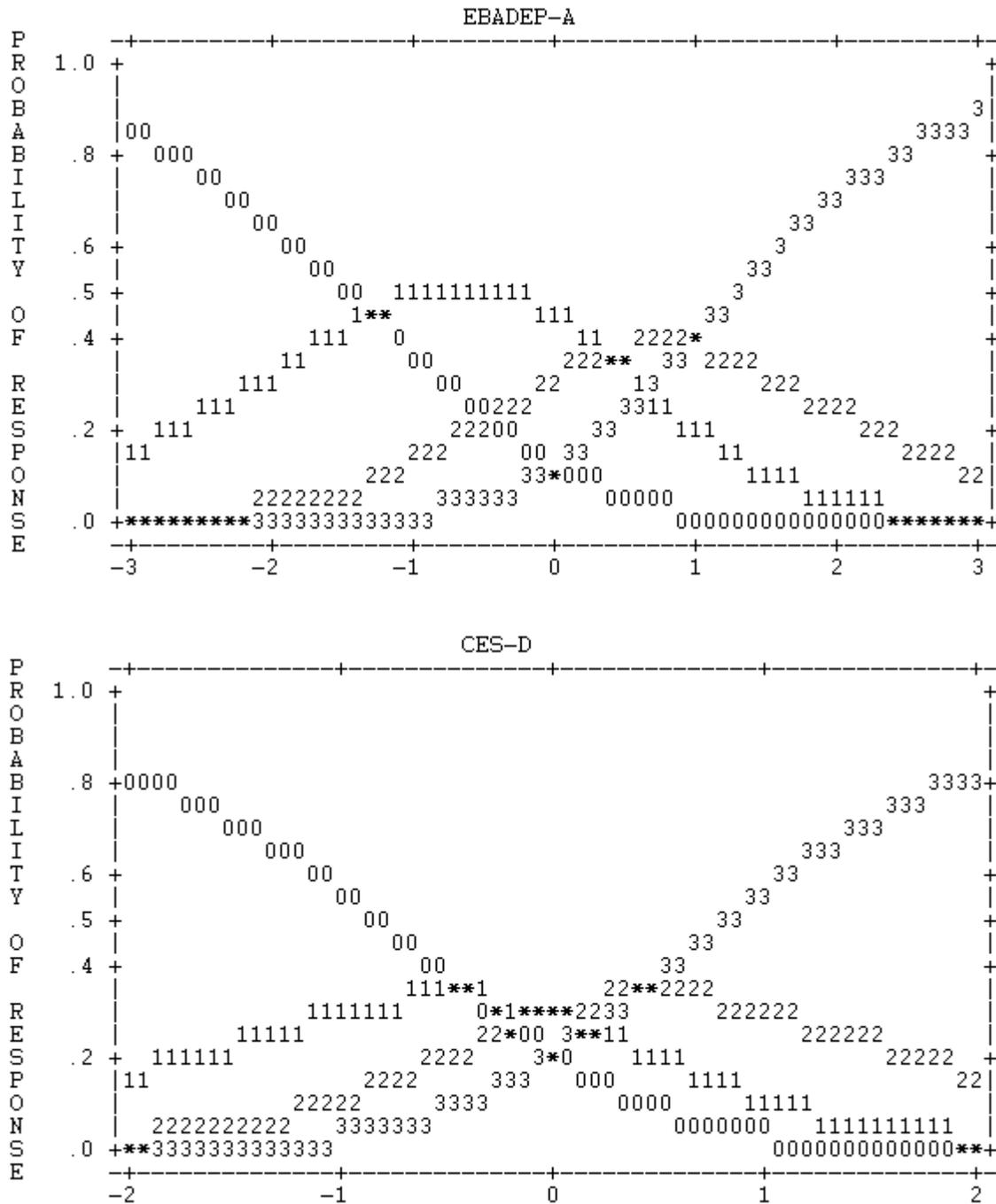


Figure 1. Distribution of response options of the instruments

Table 1
Map of Items for CES-D and EBADEP-A

| Persons | | - Items | |
|---------------|-----------|----------|-------------------------------------|
| <More> <Rare> | | | |
| 2 | + | | |
| | | | |
| | | | |
| | | EBADEP29 | |
| | . T | | |
| | | EBADEP44 | |
| | . | CESD9 | |
| 1 | . + | CESD2 | EBADEP10 |
| | . | EBADEP12 | EBADEP33 EBADEP34 |
| | # S | CESD15 | CESD3 EBADEP30 |
| | . | CESD17 | CESD19 CESD20 EBADEP21 |
| | .# T | CESD1 | EBADEP13 EBADEP28 EBADEP32 |
| | .## | EBADEP11 | EBADEP41 EBADEP42 EBADEP5 EBADEP9 |
| | .### | CESD10 | CESD6 EBADEP20 EBADEP35 EBADEP43 |
| | | EBADEP8 | |
| 0 | .## +M | CESD13 | EBADEP16 EBADEP17 EBADEP19 EBADEP22 |
| | | EBADEP3 | EBADEP31 EBADEP4 |
| | ##### | CESD18 | EBADEP18 EBADEP24 EBADEP40 |
| | .##### S | CESD11 | CESD14 EBADEP39 EBADEP6 |
| | ##### | CESD7 | EBADEP23 EBADEP25 EBADEP26 EBADEP7 |
| | .##### | CESD4 | EBADEP2 EBADEP27 EBADEP36 |
| | .##### S | CESD5 | EBADEP1 EBADEP15 EBADEP45 |
| | .##### | EBADEP14 | EBADEP37 EBADEP38 |
| -1 | .##### + | | |
| | .##### M | | |
| | .##### | | |
| | .##### T | | |
| | ##### | | |
| | .##### | CESD16 | CESD8 |
| | .##### | | |
| -2 | .##### S+ | CESD12 | |
| | #### | | |
| | .# | | |
| | .### | | |
| | . | | |
| | ## T | | |
| | . | | |
| -3 | .# + | | |
| | . | | |
| | . | | |
| | . | | |
| | . | | |
| -4 | . + | | |
| | . | | |
| | . | | |
| | . | | |
| -5 | . + | | |
| <less> <freq> | | | |

Each "#" is 4. Each "." is 1 to 3

Table 2

Conversion of scores into Thetas for CES-D

| Score | Measure | S.E. | Score | Measure | S.E. | Score | Measure | S.E. |
|-----------|--------------|------|-------|---------|------|-------|---------|------|
| 0 | -5,29E | 1,82 | 21 | -0,66 | 0,28 | 42 | 0,76 | 0,27 |
| 1 | -4,10 | 0,99 | 22 | -0,59 | 0,27 | 43 | 0,84 | 0,27 |
| 2 | -3,42 | 0,70 | 23 | -0,52 | 0,27 | 44 | 0,91 | 0,28 |
| 3 | -3,02 | 0,58 | 24 | -0,44 | 0,27 | 45 | 0,99 | 0,28 |
| 4 | -2,73 | 0,51 | 25 | -0,37 | 0,26 | 46 | 1,08 | 0,29 |
| 5 | -2,49 | 0,46 | 26 | -0,31 | 0,26 | 47 | 1,16 | 0,30 |
| 6 | -2,30 | 0,43 | 27 | -0,24 | 0,26 | 48 | 1,25 | 0,31 |
| 7 | -2,13 | 0,40 | 28 | -0,17 | 0,26 | 49 | 1,35 | 0,32 |
| 8 | -1,98 | 0,38 | 29 | -0,10 | 0,26 | 50 | 1,45 | 0,33 |
| 9 | -1,84 | 0,37 | 30 | -0,04 | 0,26 | 51 | 1,57 | 0,34 |
| 10 | -1,71 | 0,35 | 31 | 0,03 | 0,26 | 52 | 1,69 | 0,36 |
| 11 | -1,59 | 0,34 | 32 | 0,09 | 0,25 | 53 | 1,83 | 0,38 |
| 12 | -1,47 | 0,33 | 33 | 0,16 | 0,25 | 54 | 1,98 | 0,41 |
| 13 | -1,37 | 0,32 | 34 | 0,22 | 0,26 | 55 | 2,17 | 0,45 |
| 14 | -1,27 | 0,31 | 35 | 0,29 | 0,26 | 56 | 2,39 | 0,50 |
| 15 | -1,17 | 0,31 | 36 | 0,35 | 0,26 | 57 | 2,67 | 0,57 |
| 16 | -1,08 | 0,30 | 37 | 0,42 | 0,26 | 58 | 3,07 | 0,70 |
| 17 | -0,99 | 0,29 | 38 | 0,49 | 0,26 | 59 | 3,76 | 1,00 |
| 18 | -0,90 | 0,29 | 39 | 0,55 | 0,26 | 60 | 4,96E | 1,82 |
| 19 | -0,82 | 0,28 | 40 | 0,62 | 0,26 | | | |
| 20 | -0,74 | 0,28 | 41 | 0,69 | 0,27 | | | |

Table 3

Conversion of scores into Thetas for EBADEP-A

| Score | Measure | S.E. | Score | Measure | S.E. | Score | Measure | S.E. |
|-----------|--------------|------|-------|---------|------|-------|---------|------|
| 0 | -6,32E | 1,83 | 46 | -0,6 | 0,19 | 92 | 0,88 | 0,18 |
| 1 | -5,10 | 1,01 | 47 | -0,56 | 0,19 | 93 | 0,91 | 0,19 |
| 2 | -4,39 | 0,72 | 48 | -0,53 | 0,19 | 94 | 0,95 | 0,19 |
| 3 | -3,97 | 0,59 | 49 | -0,49 | 0,19 | 95 | 0,98 | 0,19 |
| 4 | -3,67 | 0,52 | 50 | -0,46 | 0,19 | 96 | 1,02 | 0,19 |
| 5 | -3,43 | 0,46 | 51 | -0,42 | 0,18 | 97 | 1,06 | 0,19 |
| 6 | -3,23 | 0,43 | 52 | -0,39 | 0,18 | 98 | 1,09 | 0,19 |
| 7 | -3,06 | 0,40 | 53 | -0,36 | 0,18 | 99 | 1,13 | 0,19 |
| 8 | -2,91 | 0,38 | 54 | -0,32 | 0,18 | 100 | 1,17 | 0,19 |
| 9 | -2,78 | 0,36 | 55 | -0,29 | 0,18 | 101 | 1,20 | 0,20 |
| 10 | -2,65 | 0,34 | 56 | -0,26 | 0,18 | 102 | 1,24 | 0,20 |
| 11 | -2,54 | 0,33 | 57 | -0,23 | 0,18 | 103 | 1,28 | 0,20 |
| 12 | -2,44 | 0,31 | 58 | -0,19 | 0,18 | 104 | 1,32 | 0,20 |
| 13 | -2,34 | 0,30 | 59 | -0,16 | 0,18 | 105 | 1,36 | 0,20 |
| 14 | -2,26 | 0,29 | 60 | -0,13 | 0,18 | 106 | 1,41 | 0,21 |
| 15 | -2,17 | 0,29 | 61 | -0,10 | 0,18 | 107 | 1,45 | 0,21 |
| 16 | -2,09 | 0,28 | 62 | -0,07 | 0,18 | 108 | 1,49 | 0,21 |
| 17 | -2,01 | 0,27 | 63 | -0,03 | 0,18 | 109 | 1,54 | 0,21 |
| 18 | -1,94 | 0,27 | 64 | 0,00 | 0,18 | 110 | 1,59 | 0,22 |
| 19 | -1,87 | 0,26 | 65 | 0,03 | 0,18 | 111 | 1,63 | 0,22 |
| 20 | -1,81 | 0,26 | 66 | 0,06 | 0,18 | 112 | 1,68 | 0,22 |
| 21 | -1,74 | 0,25 | 67 | 0,09 | 0,18 | 113 | 1,74 | 0,23 |
| 22 | -1,68 | 0,25 | 68 | 0,12 | 0,18 | 114 | 1,79 | 0,23 |
| 23 | -1,62 | 0,24 | 69 | 0,15 | 0,18 | 115 | 1,84 | 0,24 |
| 24 | -1,56 | 0,24 | 70 | 0,18 | 0,18 | 116 | 1,9 | 0,24 |
| 25 | -1,51 | 0,23 | 71 | 0,21 | 0,18 | 117 | 1,96 | 0,25 |
| 26 | -1,46 | 0,23 | 72 | 0,24 | 0,18 | 118 | 2,02 | 0,25 |
| 27 | -1,40 | 0,23 | 73 | 0,28 | 0,18 | 119 | 2,09 | 0,26 |
| 28 | -1,35 | 0,22 | 74 | 0,31 | 0,18 | 120 | 2,16 | 0,27 |
| 29 | -1,30 | 0,22 | 75 | 0,34 | 0,18 | 121 | 2,23 | 0,28 |
| 30 | -1,25 | 0,22 | 76 | 0,37 | 0,18 | 122 | 2,31 | 0,28 |
| 31 | -1,21 | 0,22 | 77 | 0,40 | 0,18 | 123 | 2,4 | 0,29 |
| 32 | -1,16 | 0,21 | 78 | 0,43 | 0,18 | 124 | 2,49 | 0,31 |
| 33 | -1,12 | 0,21 | 79 | 0,46 | 0,18 | 125 | 2,58 | 0,32 |
| 34 | -1,07 | 0,21 | 80 | 0,49 | 0,18 | 126 | 2,69 | 0,34 |
| 35 | -1,03 | 0,21 | 81 | 0,52 | 0,18 | 127 | 2,81 | 0,36 |
| 36 | -0,99 | 0,20 | 82 | 0,55 | 0,18 | 128 | 2,95 | 0,38 |
| 37 | -0,94 | 0,20 | 83 | 0,59 | 0,18 | 129 | 3,10 | 0,41 |
| 38 | -0,90 | 0,20 | 84 | 0,62 | 0,18 | 130 | 3,28 | 0,45 |

| | | | | | | | | |
|----|-------|------|----|------|------|-----|-------|------|
| 39 | -0,86 | 0,20 | 85 | 0,65 | 0,18 | 131 | 3,51 | 0,50 |
| 40 | -0,82 | 0,20 | 86 | 0,68 | 0,18 | 132 | 3,79 | 0,57 |
| 41 | -0,79 | 0,20 | 87 | 0,71 | 0,18 | 133 | 4,19 | 0,70 |
| 42 | -0,75 | 0,19 | 88 | 0,75 | 0,18 | 134 | 4,88 | 1,00 |
| 43 | -0,71 | 0,19 | 89 | 0,78 | 0,18 | 135 | 6,08E | 1,82 |
| 44 | -0,67 | 0,19 | 90 | 0,81 | 0,18 | | | |
| 45 | -0,64 | 0,19 | 91 | 0,85 | 0,18 | | | |

Table 4

Crosstabulation between the diagnostic categories of the instruments

| | | CES-D | | | |
|----------|---------------|--------|---------------|-----------|-------|
| | | | Non depressed | Depressed | Total |
| EBADEP-A | Non depressed | N | 222 | 112 | 334 |
| | | % | 82,5% | 35,0% | 56,7% |
| | Depressed | N | 47 | 208 | 255 |
| | | % | 17,5% | 65,0% | 43,3% |
| Total | N | 269 | 320 | 589 | |
| | % | 100,0% | 100,0% | 100,0% | |

Artigo 4

Tradução para o Português, Adaptação e Análise Fatorial da *Reasons for Living Scale*

Embora o suicídio, ao menos do ponto de vista epidemiológico, tenha sido amplamente estudado internacional e nacionalmente, a ideação suicida tem recebido pouco destaque em publicações. Ainda que internacionalmente possam ser encontrados vários instrumentos para sua avaliação, o mesmo não ocorre no Brasil. O objetivo do presente estudo foi realizar os procedimentos iniciais de adaptação do *Reasons for Living Scale*, a qual avalia indiretamente a ideação suicida por meio de motivos para se viver. Após procedimentos de tradução para o Português, a escala foi aplicada em 619 estudantes, com média aproximada de 24 anos (DP = 8,61), a maioria do sexo feminino (74,5%). Foram realizados três conjuntos de Análises de Componentes Principais e exploradas também as características de consistência interna da escala. A melhor organização fatorial, tanto estatística quanto semântica, para a versão brasileira do instrumento se deu com 28 itens, divididos em três fatores. Estudos posteriores são recomendados a fim de que a estrutura sugerida seja corroborada.

Palavras-chave: comportamento autodestrutivo; suicídio; consistência interna

Translation into Portuguese, Adaptation and Factor Analysis of the Reasons for Living Scale

Although the suicide, at least from an epidemiological standpoint, has been widely studied both internationally and nationally, suicidal ideation has received little attention in publications. Although internationally it can be found several instruments for evaluation, the same does not happen in Brazil. The aim of this study was the initial procedures of adaptation of the Reasons for Living Scale, which indirectly assesses suicidal ideation through reasons to live. After translation procedures into the Brazilian Portuguese, the scale was administered on 619 students with an average of approximately 24 years ($SD = 8,61$) most females (74,5%). We performed three groups of Principal Component Analysis and also explored the characteristics of internal consistency of the scale. The best factorial organization, both statistically and semantics, for the Brazilian version of the instrument was made with 28 items divided into three factors. Further studies are recommended so that the structure suggested can be corroborated.

Keywords: self-destructive behavior; suicide; internal consistency

Traducción al portugués, Adaptación y Análisis factorial del *Reasons for Living Scale*

Aunque el suicidio, al menos desde un punto de vista epidemiológico, se ha estudiado ampliamente a nivel internacional y nacional, ideación suicida ha recibido poca atención en las publicaciones. A pesar internacionalmente se pueden encontrar varias herramientas para la evaluación, lo mismo no sucede en Brasil. La finalidad del presente estudio fue los procedimientos iniciales para la adaptación del *Reasons for Living Scale*, el cual evalúa la ideación suicida indirectamente mediante la razón de vivir. Después de los procedimientos de traducción para el portugués, la escala fue administrada a 619 estudiantes, con un promedio de aproximadamente 24 años ($SD = 8,61$), en su mayoría mujeres (74,5%). Se realizaron tres juegos de análisis de componentes principales y también se estudió de las características de consistencia interna de la escala. La mejor organización factorial, tanto estadística como la semántica, de la versión brasileña del instrumento se realizó con 28 artículos divididos en tres factores. Estudios adicionales se recomiendan de modo que la estructura sugerida queda corroborada.

Palabras clave: comportamiento autodestructivo; el suicidio; la coherencia interna

Introdução

O suicídio tem como conceito oficial “ato deliberado, iniciado e levado a cabo por uma pessoa com pleno conhecimento ou expectativa de um resultado fatal” (OMS, 2001, p.80), ou seja, trata-se de um comportamento destrutivo contra si, o qual conscientemente espera-se que o resultado seja terminar com a própria vida. Quando um comportamento autodestrutivo é iniciado com propósito de morte, ou que a consequência possa ser a morte, mas não é levado a cabo devido a intervenção de terceiros ou ainda pela falha no método, por exemplo, fala-se em tentativa de suicídio (Gvion & Apter, 2012; Nock & cols., 2012).

Por sua alta incidência mundial, trata-se de um problema de saúde pública. De acordo com a Organização Mundial de Saúde (OMS), os dados globais de mortalidade indicam uma morte por suicídio a cada 40 segundos, com expectativa de aproximadamente 1,5 milhões de ocorrências em 2020 (OMS, 2001; 2006; WHO, 1999; 2000a; 2000b). Tal estimativa é equivalente a uma morte por suicídio a cada 20 segundos e uma tentativa a cada 1 ou 2 segundos (Bertolote & Fleishmann, 2002; WHO, 1999).

Murray e Lopez (1997) apresentaram uma projeção de aumento entre 4% e 7% nas taxas mundiais de suicídio, caracterizado como “lesão intencional”, com base na década de 1990 para 2020. Em estudo semelhante, com objetivo de revisão dos dados supracitados, e a inclusão de doenças não consideradas no estudo de Murray e Lopez (1997), foi realizada projeção baseada em 2002, na qual o suicídio foi apresentado com a possibilidade de se tornar a 12^a maior causa de mortalidade em todo o mundo em 2030 (Mathers & Loncar, 2006).

Contudo, para o estudo do fenômeno, não devem ser considerados somente o número bruto de suicídios, ou ainda suas estimativas. Outros fatores, relativos ao

comportamento, tanto no que se refere ao ato consumado quanto à tentativa são imprescindíveis para a sua compreensão (Bertolote & Fleishmann, 2002). Como exemplo, podem ser encontrados estudos relatando local e forma mais utilizada por homens e mulheres, dias da semana, características socioeconômicas e nível de escolarização formal, além de, no caso de estudos brasileiros, a região do país (Baptista & Borges, 2005; Botega, & cols., 2009; Faria, Victoria, Meneghel, Carvalho, & Falk, 2006; Kliemann, 2007; Kposowa & McElvain, 2006; Lovisi, & cols., 2009; Macente & Zandonade, 2011; Macente, Santos, & Zandonade, 2009; Marín-León & Barros, 2003; Meneghel, Victoria, Faria, Carvalho, & Falk, 2004; Parente, Soares, Araújo, Cavalcante, & Monteiro, 2007; dentre outros).

No que concerne especificamente às tentativas, diversos estudos podem ser citados, relativos aos diferentes grupos de idades e escolarização (Feijó, Raupp, & John, 1999; Marcondes Filho & cols., 2002; Prieto & Tavares, 2005; Rapeli & Botega, 1998), e especificamente nos estudos brasileiros, aos estados (Macente & cols., 2009; Pires, Caldas & Recena, 2005; Souza, & cols., 2011; Teixeira & Luis, 1997; Werneck, Hasselmann, Phebo, Vieira & Gomes, 2006). Entretanto, a característica mais marcante envolve a discrepância encontrada entre o número de tentativas e suicídios consumados, em relação ao gênero. Por um lado, o sexo feminino é apresentado como o grupo com maior prevalência de tentativas, ao passo que o masculino recebe destaque em relação aos atos suicidas concretizados, fato que leva a discussão sobre os possíveis motivos para esta diferença. Alguns autores afirmam que talvez as mulheres realizem um número maior de tentativas utilizando maneiras menos eficazes, enquanto os homens escolhem métodos mais violentos, o que pode influenciar no número de registros (Baptista, 2004).

Ressalta-se que os dados disponíveis não devem ser considerados como conclusivos, tanto pelo tipo de metodologia utilizada para coleta do material nas diferentes pesquisas, quanto pela dificuldade em encontrar registros mais precisos, o que gera diferenças nos índices apresentados e dificuldade no que se refere à generalização dos resultados (Baptista, 2004). Dessa forma, embora sejam de alta relevância, o estudo de dados sobre o número de suicídios consumados ou de tentativas não devem ser os únicos aspectos considerados no estudo do fenômeno suicida, mas também devem ser consideradas as investigações sobre ideação suicida, ou seja, a vulnerabilidade do indivíduo ao comportamento, apresentada na forma de planos suicidas ou pensamentos de morte.

Para Nock e colaboradores (2008), a ideação suicida é caracterizada como parte dos comportamentos suicidas, juntamente com os planos de suicídio, os quais se referem ao planejamento de um ou mais métodos específicos e da tentativa, considerada como a possibilidade de prática de um ou mais artifícios articulados. Por sua vez, para a Organização Mundial de Saúde, o comportamento suicida engloba as características desde os pensamentos, até as tentativas e o ato consumado. A ideação, especificamente, indica o desenvolvimento de ideias e pensamentos sobre o suicídio e sobre possíveis métodos, podendo ou não levar a um resultado definitivo (WHO, 2002).

De acordo com Werlang, Borges e Fensterseifer (2005), o estudo pormenorizado entre a ideação e os demais comportamentos suicidas é questão crucial para a determinação de intervenções preventivas nas diversas populações. Para tal, são importantes as pesquisas relacionadas aos fatores de risco e de proteção ao suicídio, bem como investigações que se propõem a entender o funcionamento dos comportamentos suicidas.

Isto posto, percebe-se que, dentre os comportamentos suicidas apresentados, desde a ideação, até a tentativa e o suicídio propriamente dito, a primeira é a categoria com o menor número de publicações, quando comparada às demais, talvez pela sua característica mais subjetiva, mais relacionada a pensamentos e crenças que a comportamentos observáveis. Contudo, como dito anteriormente, trata-se de um conceito também relevante para o estudo do fenômeno, sendo inclusive apresentado por alguns autores como fator de predição de comportamentos autodestrutivos (Beck, & cols., 1979; 1988; Nock, & cols., 2008; Werlang, Borges & Fensterseifer, 2005; WHO, 2002).

Percebe-se que por se tratar de um tema complexo e multifatorial, a investigação empírica do suicídio envolve os mais variados aspectos, tanto individuais, quanto sociais, culturais e também psicológicos. Entretanto, pode-se observar que há uma tendência na publicação de estudos relativos a características epidemiológicas em grupos específicos (Baptista & Borges, 2005; Botega & cols., 2009; Faria, & cols., 2006; Kliemann, 2007; Kposowa & McElvain; 2006; Lovisi, & cols., 2009; Macente & Zandonade, 2011; Macente, Santos & Zandonade, 2009; Marín-León & Barros, 2003; Mello-Santos, Bertolote, & Wang, 2005; Meneghel, & cols., 2004; Parente, & cols., 2007).

Tal diferença no número de publicações pode ser prejudicial, tanto no sentido de compreensão do fenômeno, quanto no que diz respeito ao confronto dos mitos relativos ao suicídio (Chachamovich, & cols., 2009). De fato, informações disponíveis acerca do comportamento suicida, principalmente no que diz respeito à ideação, ainda não são apresentadas em caráter conclusivo, devido a diversos fatores, os quais incluem dificuldades inerentes ao estudo do tema e também ao tipo de dados divulgados, ligados a

ideias de senso comum, não baseadas em informações provindas de resultados de pesquisas científicas (Barrero, 2005; WHO, 1999).

No que concerne à avaliação do suicídio, tanto relativo a identificação do risco quanto da ideação, várias escalas internacionais podem ser citadas, como por exemplo *Scale for Predicting Subsequent – Suicidal Behavior*, *Suicide Risk Measure*, *Suicide Ideation Scale*, *Suicide Ideation Questionnaire*, *Suicide Behavior Questionnaire*, *Reasons For Living Scale*, e *Beck Scale for Suicide Ideation* (Range & Knott, 1997). Entretanto, no que diz respeito especificamente ao Brasil, somente o Inventário de Ideação Suicida de Beck (BSI) possui tradução, adaptação e validação (Cunha & Werlang, 1996; Cunha, 2001), sendo o único instrumento aprovado para utilização clínica pelo SATEPSI (CFP, 2011).

De acordo com Kaplan e colaboradores (1994), a ideação suicida é um construto que pode ser considerado de avaliação delicada, uma vez que abrange questionamentos íntimos sobre tipos de pensamentos e crenças que as pessoas podem não se sentir a vontade em se expressar. Nestes termos, a utilização de escalas de autoavaliação pode ser vantajoso em comparação a entrevistas clínicas e dentre as diferentes possibilidades de pesquisa, apresenta-se o instrumento *Reasons for Living Scale* ([RFL] Linehan & cols., 1983).

Diferentes autores consideram o estudo dos motivos ou razões para se viver como uma alternativa para se investigar a ideação suicida (Linehan, & cols., 1983; Osman, & cols., 1993; 1999; Pergher, Stein & Wainer, 2004). O RFL é uma escala que avalia indiretamente a ideação por meio de questionamentos primordialmente relativos a diferentes motivos para manter-se vivo, ou seja, crenças e razões para que a pessoa não apresente comportamento autodestrutivo. Originalmente ela possui uma divisão em seis

fatores, quais sejam, “Crenças de sobrevivência e enfrentamento”, “Responsabilidade para com a família”, “Interesses relacionados às crianças (filhos)”, “Medo do suicídio”, “Medo da desaprovação social”, e “Objeções morais” (Linehan & cols., 1983).

As qualidades psicométricas da RFL foram verificadas primeiramente em sua construção, por meio de dois estudos iniciais (Linehan & cols., 1983). O Primeiro contou com 431 participantes, divididos em duas amostras, de 218 pessoas (96 homens e 112 mulheres) e 213 (124 homens e 89 mulheres). Foi realizada análise dos componentes principais, com rotação ortogonal *Varimax*, em cada amostra. No total, foram rodadas quatro análises fatoriais, as quais conduziram a resultados muito similares. A partir daí, a escala foi dividida nos seis fatores supracitados. A consistência interna foi apurada pelo alfa de Cronbach, sendo encontrados valores entre 0,72 e 0,89. Posteriormente, no segundo estudo, participaram 63 homens e 112 mulheres pacientes de um Hospital Universitário, os quais responderam, além da RFL, um Questionário de Comportamentos Suicidas (QCS). As correlações entre os fatores obtidos pela RFL e os componentes do QCS confirmaram as evidências de validade convergente.

Uma terceira pesquisa foi realizada com objetivo de validação da escala, em uma amostra adulta composta por 205 pacientes psiquiátricos (Osman & cols., 1999). Do total, 105 eram homens, com média de idade igual a 31,4 anos (DP = 7,7), 31 deles com histórico de tentativa de suicídio, e 100 mulheres, com idade média igual a 32,2 anos (DP = 7,9), 44 com histórico de tentativa de suicídio. O alfa de Cronbach encontrado para a escala total foi de 0,93, enquanto os índices dos fatores variaram entre 0,77 (Objeções Morais) e 0,95 (Crenças de sobrevivência e enfrentamento). Pela análise fatorial confirmatória, o modelo com seis fatores obteve melhor ajuste que o modelo unifatorial, com 79% da covariância

explicada. Foram também verificadas as cargas fatoriais, as quais variaram entre 0,37 e 0,93, juntamente com as correlações entre os fatores, entre 0,24 e 0,57.

Embora a quantidade de instrumentos disponíveis mundialmente possa ser considerada suficiente, observa-se que no Brasil o mesmo não ocorre. Além disso, paradoxalmente, um número relativamente pequeno de pesquisas procuram analisar como são realizadas as avaliações (Chachamovich, & cols., 2009). A Resolução nº002/2003 do Conselho Federal de Psicologia (CFP, 2003), com base na Comissão Internacional de Testes (ITC, 2000) e nos *Standards for Educational and Psychological Test (American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME], 1999)* regulamenta a utilização clínica de testes psicológicos, relacionando as qualidades psicométricas básicas para que um questionário, escala ou inventário possam ser utilizados, especificamente a precisão e a validade (Noronha & Alchieri, 2004; Noronha & Vendramini, 2003; Pasquali, 1999; Urbina, 2004).

De fato, toda a estrutura histórica envolvendo a Avaliação Psicológica permitiu que se pudesse chegar à discussão da necessidade da construção de instrumentos confiáveis, com qualidades reconhecidas internacionalmente, de forma que há um grande foco centrado na construção e validação de testes (Cruz, 2002; Noronha, 2009). Entretanto, deve-se considerar que não somente a construção deve receber atenção e cuidados técnicos, mas também sua revisão (CFP, 2003; Noronha & Alchieri, 2002).

A revisão de um instrumento, de acordo com Noronha e Alchieri (2002) é ainda pouco discutida no Brasil, mas muito importante na medida em que nenhum teste deve permanecer cristalizado ao longo dos anos, enquanto o contexto de sua construção pode

sofrer modificações e o construto, atualizações em sua conceituação. Este é um dos motivos pelos quais se fazem necessários diferentes estudos ao se introduzir à população brasileira um instrumento oriundo de outro país, afinal, não somente as diferenças de linguagem estão envolvidas, mas também culturais e contextuais (Manzi-Oliveira, Balarini, Marques, & Pasian, 2011).

Assim, não se pode considerar que por um teste apresentar bons índices de confiabilidade e diferentes evidências de validade em seu país original, que a mesma estrutura e as mesmas propriedades serão visualizadas após uma tradução (AERA, APA, & NCME, 1999; Manzi-Oliveira, & cols., 2011). Por isso, antes da utilização de um teste, independente se ter sido construído no país ou traduzido e adaptado, ele precisa ter a sua validade estudada, para que sua aplicação posterior seja possível em pessoas com as características semelhantes (Manzi-Oliveira, & cols., 2011; Urbina, 2004).

O presente estudo se situa dentro deste contexto de adaptação e revisão de um instrumento psicológico. A partir da possibilidade de se avaliar a ideação suicida por meio de uma perspectiva cognitiva, dentro de um sistema de crenças pessoais sobre motivos para não se cometer o suicídio (Linehan, & cols., 1983; Osman, & cols., 1993; 1999), este trabalho teve como objetivo principal realizar a tradução e adaptação da *Reasons for Living Scale* (RFL) para o português do Brasil, além de buscar evidências de validade de conteúdo e baseadas na estrutura interna.

Método

Participantes

Participaram deste estudo 619 estudantes, a maioria universitários (95,5%) de seis instituições de ensino superior de regiões diferentes em Minas Gerais e São Paulo, com idade média de 23,68 anos (DP = 8,61), variando entre 17 e 63 anos. Os 28 participantes não universitários (4,5%) faziam parte de um curso profissionalizante e de ensino de língua estrangeira. A amostra foi composta majoritariamente por mulheres (N = 461; 74,5%) e solteiros (N = 509; 82,2%).

Instrumentos

1. Questionário de identificação

Este questionário, elaborado pelos membros do Laboratório em Avaliação Psicológica em Saúde Mental III (Lapsam 3) da Universidade São Francisco, foi utilizado com o objetivo de caracterização da amostra e coleta de informações relevantes para o estudo, às quais se destacam sexo e idade, região de moradia e estado civil dos participantes.

2. Escala de Razões para Viver (ERV/RFL)

A Escala de Razões para Viver (ERV), do Inglês *Reasons for Living Scale* (RFL), foi construída com a intenção de se avaliar a ideação suicida de forma indireta, contendo questões relativas às crenças sobre a importância em se estar vivo. Ela é composta originalmente por 48 itens, em formato *Likert* de seis pontos, os quais variam entre “sem importância” a “extremamente importante”. Para cada item é requisitado ao participante

que seja feita uma análise de seu conteúdo com base na importância que a sentença apresentada tem para que ele não se mate (Linehan & cols., 1993).

O instrumento passou por estudos de suas qualidades psicométricas, primeiramente durante sua construção e posteriormente em amostras clínicas e não clínicas (Linehan & cols., 1993; Osman, & cols., 1993; 1999). Ele foi apresentado originalmente por Linehan e colaboradores (1993) contendo seis fatores, os quais são calculados com base nas médias de pontuação de seus respectivos grupos de itens. São eles: “Crenças de sobrevivência e enfrentamento” (*Survival Coping Beliefs* – SCB, 24 itens); “Responsabilidade para com a família” (*Responsibility to Family* – RF, 7 itens); “Interesses relacionados às crianças (filhos)” (*Child-Related Concerns* – CRC, 3 itens); “Medo do suicídio (*Fear of Suicide* – FS, 7 itens)”; “Medo da desaprovação social” (*Fear of Social Disapproval* – FSD, 3 itens); e “Objeções morais” (*Moral Objections* – MO, 4 itens).

Procedimentos

O procedimento de tradução da escala para o Português brasileiro ocorreu após autorização da autora, Profa Dra Marsha Linehan. O instrumento passou inicialmente por um procedimento de Tradução e posteriormente pela Retro-tradução por dois diferentes tradutores-técnicos, bilíngues em Português e Inglês, a fim de que os significados e conceitos apresentados não fossem deturpados, no caso de um eventual erro de tradução. A versão traduzida foi comparada à original, revisada quanto ao seu conteúdo semântico e aplicada, em estudo piloto, em 35 estudantes universitários, 21 (60%) do estado da Bahia e 14 (40%) de São Paulo. Foi requisitado aos participantes que eles anotassem ou comentassem sobre a possível dificuldade de compreensão da instrução ou de algum dos

itens. Não foi entregue nenhuma consideração por escrito em relação a este ponto, entretanto pequenas mudanças em relação à instrução foram sugeridas (Baptista & Gomes, 2012).

A média de pontuação da amostra piloto foi de 207,77 pontos (DP = 24,69), variando entre 151 a 264. As pontuações médias de cada fator foram comparadas com as da amostra original pelo método de teste *t One Sample*. Mesmo com um número consideravelmente menor de participantes, foram encontrados valores estatisticamente mais altos para a amostra brasileira nos fatores “Interesses relacionados às crianças (filhos)”, “Medo do suicídio” e “Objeções morais”. Os autores discutiram sobre as influências culturais em relação ao estudo da ideação suicida, sendo que as médias não significativas poderiam indicar que, culturalmente falando, as crenças sobre enfrentamento, responsabilidade com a família e aprovação ou desaprovação social podem possuir características parecidas em ambos os contextos (Baptista & Gomes, 2012).

Após a finalização do processo de tradução da ERV, puderam ser realizados os procedimentos do presente estudo. Inicialmente, foram considerados todos os cuidados éticos, sendo aplicados os instrumentos após aprovação do Comitê de Ética em Pesquisa da Universidade São Francisco (Protocolo CAE 0422.0.142.000-11). Inicialmente foram contabilizados 640 protocolos, entretanto foram desconsiderados 21 nos quais a escala não foi respondida completamente.

Resultados

Primeiramente, foi verificada a existência de *missing*, ou seja, questões as quais os participantes não responderam. Os protocolos com até 3 *missing* foram mantidos para

análise, havendo a substituição pelos valores médios individuais. Entretanto, é interessante ressaltar que o item 48 (Eu não gostaria que minha família pensasse que eu fui covarde ou um egoísta) foi deixado em branco por 23 participantes (3,7%), enquanto o item 28 (O efeito sobre os meus filhos pode ser prejudicial) obteve quatro *missing* (0,6%). A fim de se cumprir os objetivos do presente estudo, os resultados foram separados em 3 estudos.

Estudo 1

Inicialmente foi realizada uma Análise de Componentes Principais (ACP), de caráter exploratório, com rotação *Varimax* e supressão das cargas fatoriais menores que 0,50, assim como realizado inicialmente por Linehan e colaboradores (1983). O teste de esfericidade de Bartlett ($X^2 = 10646,48$, $gl = 1128$; $p = 0,000$) e a medida de adequação Kaiser-Meyer-Olkin ($KMO = 0,90$), indicaram alta possibilidade de separação da escala em fatores. O teste de confiabilidade, por meio do alfa de Cronbach gerou um índice de 0,90 para os 48 itens. A comunalidade variou entre 0,36 (itens 4 e 25) e 0,74 (item 21).

Foram gerados inicialmente 11 fatores, os quais explicaram 57,56% da variância total, embora a análise do gráfico *Scree Plot* tenha sugerido a presença de três fatores, considerando os *eigenvalues* acima de 1. Baseado na leitura do *Scree plot*, foi realizada nova ACP, seguindo os mesmos critérios que a anterior, sugerindo a organização em três fatores, os quais explicaram 34,80% da variância total.

Como as cargas fatoriais abaixo de 0,50 sugeriram a exclusão de itens, foram realizadas outras duas análises até que não fosse sugerida a exclusão de mais nenhuma variável. Com a retirada de 20 itens, foram verificados os testes de adequação ($KMO = 0,892$) e de esfericidade de Bartlett ($X^2 = 6180,21$; $gl = 378$; $p = 0,000$), com

comunalidades variando entre 0,34 (item 19) e 0,61 (item 20) e variância total explicada de 44,46% com uma divisão em três fatores (Tabela 1).

Inserir Tabela 1 aproximadamente aqui

Após rotação, o primeiro fator agrupou 13 itens ($\alpha = 0,88$) do original “Crenças de sobrevivência e enfrentamento”, enquanto o segundo fator, com 10 itens ($\alpha = 0,84$) continham os três itens do fator “Medo da desaprovação social” e também “Medo do suicídio” (Tabela 1). O terceiro fator, por sua vez, agrupou cinco itens ($\alpha = 0,69$), dois deles do fator “Interesses relacionados às crianças (filhos)”, dois de “Responsabilidade para com a família” e um de “Objeções morais”. O alfa de Cronbach para os 28 itens resultantes desta análise foi de 0,86 e para esta estrutura foi observado 140 (37%) de resíduos não redundantes com valores absolutos maiores que 0,05 (Hair Jr, Anderson, Tatham, & Black, 2005).

Estudo 2

Embora os procedimentos originais da escala tenham sido realizadas admitindo cargas fatoriais mínimas de 0,50 para a escolha de itens, foi desenvolvida uma segunda bateria de ACPs, assumindo saturação mínima de 0,35 correspondendo a um pouco mais de 12% de contribuição de cada item para o fator. Para as Ciências Sociais, este pode ser considerado um bom parâmetro, uma vez que devido a amostra ser de aproximadamente 600 pessoas, as cargas possuem significância mínima de 0,05 e nível de poder de 80% (Hair Jr, & cols., 2005).

A partir da leitura do gráfico *Scree Plot* e assumindo minimamente autovalores acima de 1, foram realizadas outras três ACPs com a extração de 3 fatores, os quais passaram a explicar 38,45% da variância total, após exclusão dos itens sugeridos. Foram verificados os índices de adequação ($KMO = 0,903$) e de esfericidade ($X^2 = 9448$; $gl = 820$; $p = 0,000$) e as comunalidades variaram entre 0,18 (item 15) e 0,60 (item 20). Observou-se, porém, que três itens apresentaram cargas fatoriais significativas para mais de um fator.

Quando um item apresenta mais de uma carga significativa, é recomendada a escolha para o fator com maior saturação. Entretanto, este procedimento pode dificultar a interpretação dos fatores, uma vez que embora o item seja selecionado para um fator específico, é requisitado ao pesquisador que ele inclua a variável em ambos os blocos de itens, teoricamente. Como o objetivo do procedimento é minimizar o número de cargas fatoriais de forma que cada linha da matriz rotacionada apresente uma variável com uma carga associada a um único fator (Hair Jr, & cols., 2005), foram realizadas novas ACPs, com rotação *Varimax*.

Para esta organização com três fatores, não foram aceitos itens que apresentaram cargas fatoriais abaixo de 0,35, bem como os itens com mais de uma carga fatorial significativa, resultando em uma estrutura de 32 itens com alfa de Cronbach de 0,87. O fator 1 agrupou 17 itens ($\alpha = 0,87$) correspondentes a “Crenças de sobrevivência e enfrentamento”, o fator 2 congregou 12 itens ($\alpha = 0,83$) referentes a “Medo do suicídio”, “Medo da desaprovação social” e “Objeções Morais”, e o fator 3 ficou com 3 itens ($\alpha = 0,71$), sobre “Responsabilidade para com a família”.

Pode-se perceber que tanto teórica quanto estatisticamente, a estrutura resultante apresentou boa disposição dos itens. Entretanto, a presença de somente três itens em um

fator pode não ser uma organização equânime, visto que os demais fatores apresentaram um número consideravelmente maior. Com isso, um novo conjunto de ACPs foi realizado, forçando a divisão da escala em dois fatores, com nível de saturação mínima de 0,35. Após rotação e exclusão de itens, foi sugerida uma estrutura de 44 itens com $KMO = 0,90$ ($X^2 = 9623,29$; $gl = 946$; $p = 0,000$), com alfa de Cronbach igual a 0,90.

As comunalidades variaram entre 0,13 (item 1) e 0,59 (item 20) e os fatores explicaram 32,12% da variância total, com 340 (35%) de resíduos não redundantes, com valores absolutos maiores que 0,05. O primeiro fator congregou 27 itens de “Crenças de sobrevivência e enfrentamento”, “Responsabilidade para com a família”, “Interesses relacionados às crianças (filhos)” e “Objeções morais” ($\alpha = 0,90$) e o segundo fator agrupou 17 itens de “Responsabilidade para com a família”, “Interesses relacionados às crianças (filhos)”, “Medo do suicídio”, “Medo da Desaprovação social” e “Objeções morais” ($\alpha = 0,86$).

Estudo 3

Finalizando as análises exploratórias, foram realizadas também análises usando critério *a priori* (Hair Jr, & cols., 2005), com base nos dados descritos por Linehan e colaboradores (1983), ou seja, compelindo a divisão de seis fatores para a versão traduzida da RFL. Primeiramente foram verificados o teste de esfericidade de Bartlett ($X^2 = 10646,48$; $gl = 1128$; $p = 0,000$) e o de adequação da amostra ($KMO = 0,901$). A comunalidade variou entre 0,09 (item 4) e 0,69 (item 47) e os seis fatores explicaram 45,43% da variância total. Uma vez que a supressão de itens com carga fatorial menor que 0,50 sugeriu a exclusão de alguns itens, foram realizadas outras duas análises resultando em uma

configuração de 28 itens, com $KMO = 0,88$ ($X^2 = 5919,75$; $gl = 378$; $p = 0,000$), comunalidades variando entre 0,40 (item 45) e 0,79 (item 47), 56,62% de variância total explicada e alfa de Cronbach igual a 0,86.

Foram observadas semelhanças e diferenças em relação à configuração original determinada no estudo de Linehan e colaboradores (1983). O modelo apresentou 86 (22%) de valores residuais não redundantes acima de 0,05, sendo que o primeiro fator ficou composto por 13 itens ($\alpha = 0,88$), todos presentes no fator “Crenças de sobrevivência e enfrentamento”, enquanto o segundo, com 5 itens ($\alpha = 0,79$), agrupou todos os itens do fator “Medo da desaprovação social”, e também itens de “Medo do suicídio”. Por sua vez, os fatores 3 e 4, cada um com três itens cada ($\alpha = 0,79$ e $\alpha = 0,71$ respectivamente) apresentaram os itens do fator original “Responsabilidade para com a família”, enquanto o fator 5, com dois itens ($\alpha = 0,60$) apresentaram itens de “Medo do suicídio”, e o fator 6, também com dois itens ($\alpha = 0,62$), continha itens de “Objeções morais”.

Discussão

O presente estudo teve como objetivo apresentar uma alternativa de Avaliação Psicológica no que diz respeito à ideação suicida, utilizando uma mensuração indireta do construto, por meio da Escala de Razões para Viver. De fato, o suicídio se trata de um problema de saúde pública (Bertolote & Fleishmann, 2002; Mathers & Loncar, 2006; Murray & Lopez, 1997; OMS, 2001; 2006; WHO, 1999) e muito tem se estudado sobre sua epidemiologia, as diferenças entre gênero e idade, locais, métodos, e até os dias da semana (Baptista & Borges, 2005; Botega, & cols., 2009; Faria, & cols., 2006; Lovisi, & cols.,

2009; Macente & Zaddonade, 2011; Marín-León & Barros, 2003; Meneghel, & cols., 2004; Parente, & cols., 2007; dentre outros).

Como visto, a divulgação de estudos com intuito de investigação do comportamento suicida, embora complexa, tem sido ampliada no Brasil, mesmo que os dados encontrados não possam ser considerados conclusivos, o que pode ocorrer tanto pela dificuldade em se encontrar registros válidos quanto pelas diferenças metodológicas nos diversos estudos (Baptista, 2004). Entretanto, quando se trata do estudo da ideação suicida, um dos comportamentos cruciais para a determinação de intervenções preventivas (Werlang, Borges, & Fensterseifer, 2005), os pesquisadores encontram dificuldades, uma vez que a maior parte dos instrumentos sobre este construto específico é de outros países e somente um teste recebeu aprovação pelo SATEPSI (CFP, 2011).

Além disso, duas importantes características devem ser ressaltadas, primeiramente por se tratar de um construto de delicada medição, e envolver um assunto em que as pessoas culturalmente apresentam dificuldade em lidar e se expressar (Kaplan, & cols., 1994). Em segundo lugar a maioria dos instrumentos divulgados internacionalmente, bem como o teste brasileiro devidamente validado para uso, possuem itens os quais questionam sobre pensamentos autodestrutivos de maneira direta (Baptista, 2004).

A Escala de Razões para Viver, construída em 1983 nos Estados Unidos, apresenta-se como uma solução metodológica a estas questões. Primeiramente, por ser auto administrado, não é necessário que a pessoa verbalize suas opiniões, e em segundo lugar, por apresentar itens aos quais é requisitado que sejam feitas reflexões sobre os motivos para se estar vivo, já que sua mensuração da ideação suicida é de forma indireta (Linehan & cols., 1993; Osman, & cols., 1993; 1999).

Após os procedimentos de tradução da escala para o Português (Baptista & Gomes, 2012), foram realizadas diferentes coletas de dados, em cidades distintas do interior de São Paulo e Minas Gerais, no sentido de adaptar o instrumento para o Brasil. Ao todo três estudos foram executados, nos quais se exploraram formas de organização da escala, com diferentes números de fatores e também com distintos itens em cada um, como consequência tanto da divisão por si só, quanto dos valores mínimos admitidos para as cargas fatoriais.

Cabe ressaltar que durante uma Análise Fatorial, no caso, a realização de diferentes Análises de Componentes Principais (*principal axis factoring*) para comparação de resultados, a escolha da melhor organização, assim como a definição dos fatores não deve ser realizada de forma a considerar puramente os resultados das fórmulas estatísticas, mas também os aspectos teóricos, os resultados de estudos anteriores e a congruência semântica entre fatores e itens (Hair Jr, & cols., 2005). Dessa forma, embora matematicamente possa se dizer que o estudo 3 tenha resultado em uma melhor organização fatorial, por apresentar maior porcentagem de variância explicada e menor porcentagem em resíduos não redundantes, o Estudo 1 apresentou a melhor e mais congruente distribuição de fatores, além do respaldo semântico para os agrupamentos encontrados.

Nesta distribuição, na qual a Escala de Razões para Viver se constituiu de 28 itens, agrupados em 3 fatores, não foram percebidos itens com mais de uma carga fatorial significativa, ou seja, a mesma variável em mais de um fator. Para a escolha dos itens nos fatores assumiram-se minimamente cargas fatoriais 0,50, isto é, a participação de ao menos 25% de cada item no fator. Além disso, as variâncias explicadas por cada fator e para a

escala total foram mais que satisfatórios para a sua utilização, mesmo havendo uma porcentagem mais alta de resíduos, quando comparados aos resultados dos outros estudos.

Ao primeiro fator foi mantido o nome original “Crenças de sobrevivência e enfrentamento”, o qual agrupou itens relacionados às crenças sobre si mesmo e sobre o futuro como motivos para não se matar. Por sua vez, o segundo fator, nomeado como “Medo da morte e de desaprovação social”, contendo questões relativas aos pensamentos intimamente direcionados ao ato suicida. Por fim, o terceiro fator foi denominado de “Motivação de controle externo”, por ter agrupado os motivos externos como razões para não se tirar a própria vida, sendo quatro relacionados à família e filhos, e um relacionado a Deus.

Entretanto, faz-se importante observar que a organização da escala em seis fatores, assumindo o critério *a priori* (Hair Jr, & cols., 2005), ou seja, admitindo a organização previamente estabelecida em outros estudos, especificamente o de Linehan e colaboradores (1983) também se apresentou de forma congruente semanticamente, sendo que a escala brasileira apresentou itens diferentes em cada um dos fatores, assim como ocorreu no estudo de Osman e colaboradores (1999). Os índices de confiabilidade para cada fator e as porcentagens de variância explicada tanto para cada fator quanto para a escala como um todo podem ser considerados satisfatórios para se aceitar a estrutura proposta.

Contudo, é necessário ponderar que ambas as estruturas encontradas possuem limitações. Por um lado, com a exclusão de um número considerável de itens, uma versão reduzida do instrumento pode não conter questões importantes para um estudo minucioso de um paciente, por exemplo. Por outro lado, devem ser consideradas todas as questões

culturais envolvidas na adaptação de um instrumento, já que estas diferenças também são refletidas nos itens, o que torna a retirada de itens um aspecto positivo.

Assim, pode-se concluir que o presente estudo cumpriu, de acordo com o requisitado pelo Conselho Federal de Psicologia (CFP, 2003), com base na Comissão Internacional de Testes (ITC, 2000), o propósito de apresentar as qualidades psicométricas básicas iniciais para a Escala de Razões para Viver, versão traduzida da *Reasons for Living Scale* para o Brasil, de forma que ela pode ser caracterizada como uma versão *short-form* da mesma. À versão do instrumento composta por três fatores foram aferidas evidências de validade de conteúdo e baseadas na estrutura interna (AERA, APA, & NCME, 1999). Obviamente, estudos posteriores com os instrumentos gerados desta pesquisa se fazem necessários, tanto para dar continuidade à adaptação da escala para o Brasil quanto para confirmar as estruturas sugeridas.

Referências

- American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME]. (1999). *Standards for Psychology and Educational Testing*. Washington, DC: American Psychology Association.
- Baptista, M. N. (2004). *Suicídio e Depressão: Atualizações*. São Paulo: Guanabara Koogan.
- Baptista, M. N., & Borges, A. (2005). Suicídio: aspectos epidemiológicos em Limeira e adjacências no período de 1998 a 2002. *Estudos de Psicologia: Campinas*, 22, 425-431.
- Baptista, M. N., & Gomes, J. O. (2012). *Procedimento de tradução e estudo piloto de adaptação da Reasons for Living Scale*. Em Anais de Evento: I Congresso Internacional de Psicologia, V Congresso de Psicologia da Zona da Mata e Vertentes e o IX Encontro Juizforano de Psicologia. Apresentação Oral.
- Barrero, S. A. P. (2005). Los Mitos sobre el suicidio. La importancia de conocerlos. *Revista Colombiana de Psiquiatría*, 24(3), 386-394.
- Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of Suicidal Intention: The Scale for Suicide Ideation. *Journal of Consulting and Clinical Psychology*, 47(2), 343-352.
- Bertolote, J. M., & Fleishmann, A. (2002). A Global perspective in the epidemiology of suicide. *Suicidology*, 7(2), 6-8.
- Botega, N. J., Marín-Leó, L., Oliveira, H. B., Barros, M. B. A., Silva, V. F., & Dalgalarondo, P. (2009). Prevalência de ideação, plano e tentativa de suicídio: um

- inquérito de base populacional em Campinas, São Paulo, Brasil. *Cadernos de Saúde Pública*, 25(12), 2632-2638.
- Chachamovich, E., Stefanello, S., Botega, N., & Turecki, G. (2009). Quais são os recentes achados clínicos sobre a associação entre depressão e suicídio? *Revista Brasileira de Psiquiatria*, 31(supl. I), 618-625.
- Conselho Federal de Psicologia – CFP (2003). Resolução nº 002/2003 [On-line]. Recuperado de http://www.pol.org.br/pol/export/sites/default/pol/legislacao/legislacaoDocumentos/resolucao2003_02.pdf
- Conselho Federal de Psicologia (CFP). (2011). Sistema de Avaliação de Testes Psicológicos (SATEPSI). Disponível em: <http://www.pol.org.br/satepsi/>
- Cruz, R. M. (2002). O processo de conhecer em avaliação psicológica. In: Roberto Moraes Cruz, João Carlos Alchieri & Jamir J. Sardá Jr (Orgs.). *Avaliação e Medidas Psicológicas*. São Paulo: Casa do Psicólogo.
- Cunha, J. A. (2001). *Manual da versão em português das Escalas de Beck*. Casa do Psicólogo, São Paulo.
- Cunha, J. A., & Werlang, B. C. (1996). Um estudo com a Escala de Desesperança de Beck em grupos clínicos e não-clínicos. *Psico* 27, 189-197.
- Faria, N. M. X., Victoria, C. G., Meneghel, S. N., Carvalho, L. A., & Falk, J. V. (2006). Suicide rates in the State of Rio Grande do Sul, Brazil: association with socioeconomic, cultural and agricultural factors. *Cadernos de Saúde Pública*, 22(12), 2611-2621.

- Feijó, R. B., Raupp, A. P. G., & John, A. B. (1999). Eventos estressores de vida e sua relação com tentativas de suicídio em adolescentes. *Jornal Brasileiro de Psiquiatria*, 48(4), 151-157.
- Gvion, Y., & Apter, A. (2012). Suicide and Suicidal Behavior. *Public Health reviews*, 34(2), 92-111.
- Hair Jr., J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). Análise Multivariada de Dados. Porto Alegre: Bookman.
- International Test Commission [ITC] (2000). Guidelines on Adapting Test. *International Test Commission*. [On-Line]. Recuperado de <http://www.intestcom.org/upload/sitefiles/40.pdf>
- Kaplan, M. L., Asnis, G. M., Sanderson, W. C., Keswani, L., Lecuosa, J. M., & Joseph, S. (1994). Suicide assessment: clinical interview vs. self-report. *Journal of Clinical Psychology*, 50(2), 294-298.
- Kliemann, D. V. O. (2007). Estudo epidemiológico de óbitos por suicídio na região da grande Florianópolis de 1991 a 2005. Trabalho de Conclusão de Curso. Universidade Federal de Santa Catarina. [On-Line]. Recuperado de <http://www.bibliomed.ccs.ufsc.br/CM0614.pdf>.
- Kposowa, A. J., & McElvain, J. P. (2006). Gender, place and method of suicide. *Society of Psychiatry Epidemiology*, 41, 435-443.
- Linehan, M. M., Goodstein, J. L., Nielsen, S. L., & Chiles, J. A. (1983). Reasons for Staying Alive. When you are thinking of killing yourself. The Reasons for Living Inventory. *Journal of Consulting and Clinical Psychology*, 51(2), 276-286.

- Lovisi, G. M., Santos, S. A., Legay, L., Abelha, L., & Valencia, L. (2009). Análise epidemiológica do suicídio no Brasil entre 1980 e 2006. *Revista Brasileira de Psiquiatria*, 31(2), 586-594.
- Macente, L. B. M., Santos, E. G., & Zandonade, E. (2009) Tentativas de suicídio em município de cultura Pomerana no interior do estado do Espírito Santo. *Jornal Brasileiro de Psiquiatria*, 58(4), 238-244.
- Macente, L. B., & Zandonade, E. (2011). Estudo da série histórica de mortalidade por suicídio no Espírito Santo (de 1980 a 2006). *Jornal Brasileiro de Psiquiatria*, 60(3), 151-157.
- Manzi-Oliveira, A. B., Balarini, F. B., Marques, L. A. S., & Paisan, S. R. (2011). Adaptação cultural de instrumentos de Avaliação Psicológica: levantamento de estudos realizados no Brasil de 2000 a 2010. *Psico-USF*, 16(3), 367-381.
- Marcondes Filho, W., Mezzaroba, L., Turini, C. A., Koike, A., Motomatsu Junior, A., Shibayama, E. E. M., & Fenner, F. L. S. (2002). Tentativas de suicídio por substâncias químicas na adolescência e juventude. *Adolescencia Latinoamericana*, 3(2).
- Marín-León, L., & Barros, M. B. A. (2003). Mortes por suicídio: diferenças entre gênero e nível socioeconômico. *Revista de Saúde Pública*, 37(3), 357-363.
- Mathers, C. D., & Loncar, D. (2006). Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLoS Medicine*, 3(11), 2011-2030.
- Mello-Santos, C., Bertolote, J. M., & Wang, Y-P. (2005). Epidemiology of suicide in Brazil (1980-2000): characterization of age and gender rates of suicide. *Revista Brasileira de Psiquiatria*, 27(2), 131-134.

- Meneghel, S. N., Victoria, C. G., Faria, N. M. X., Carvalho, L. A. C., & Falk, J. W. (2004). Características epidemiológicas do suicídio no Rio Grande do Sul. *Revista de Saúde Pública, 38*(6), 804-810.
- Murray, C. J. L., & Lopez, A. D. (1997). Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *The Lancet, 349*, 1498-1997.
- Nock, M. K., Borges, G., Bromet, E. J., Cha, C. B., Kessler, R. C., & Lee, S. (2008). Suicide and Suicidal Behavior. *Epidemiologic Reviews, 30*, 133-154.
- Nock, M. K., Borges, G., Bromet, E. J., Cha, C. B., Kessler, R. C., & Lee, S. (2012). The epidemiology of suicide and suicidal behavior. Em M. K. Nock, G. Borges, & Y. Ono. *Suicide. Global Perspectives from the WHO World Mental Health Surveys*. United Kingdom: Cambridge University.
- Noronha, A. P. P. (2009). Testes Psicológicos: conceito, uso e formação do psicólogo. Em Claudio S. Hutz (org). *Avanços e polêmicas em avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Noronha, A. P. P., & Alchieri, J. C. (2002). Reflexões sobre os Instrumentos de Avaliação Psicológica. Em Ricardo Primi (org.). *Temas em avaliação psicológica*. Campinas: Imprensa Digital do Brasil
- Noronha, A. P. P., & Alchieri, J. C. (2004). Conhecimento em Avaliação Psicológica. *Estudos de Psicologia, 21*(1), 43-52.
- Noronha, A. P. P., & Vendramini, C. M. M. (2003). Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica, 16*(1), 177-182.

- Organização Mundial da Saúde – OMS (2001). *Relatório Mundial de Saúde*. Saúde Mental: nova concepção, nova esperança. [On-Line]. Recuperado de http://www.who.int/whr/2001/en/whr01_djmessage_po.pdf
- Organização Mundial de Saúde- OMS (2006). *Prevenção do Suicídio– um recurso para conselheiros*. Genebra, Departamento de Saúde Mental e de Abuso de Substâncias.
- Osman, A., Giffords, J., Jones, T., Lickiss, L., Osman, J., & Wenzel, R. (1993). Psychometric Evaluation of the Reasons for Living Inventory. *Psychological Assessment*, 5(2), 154-158.
- Osman, A., Kopper, B. A., Linehan, M. M., Barrios, F. X., Gutierrez, P. M., & Bagge, C. L. (1999). Validation of the Adult Suicidal Ideation Questionnaire and the Reasons for Living Inventory in an Adult Psychiatric Inpatient Sample. *Psychological Assessment*, 11(2), 115-123.
- Parente, A. C. M., Soares, R. B., Araújo, A. R. F., Cavalcante, I. S., Monteiro, C. F. S. (2007). Caracterização dos casos de suicídio em uma capital do nordeste brasileiro. *Revista Brasileira de Enfermagem*, 60(4).
- Pasquali, L. (1999). *Instrumentos psicológicos: manual prático de elaboração*. Brasília: LabPAM & IBAPP.
- Pergher, G. K., Stein, L. M., & Wainer, R. (2004). Estudos sobre a memória na depressão: achados e implicações para a terapia cognitiva. *Revista de Psiquiatria Clínica*, 31(2), 82-90.
- Pires, D. X., Caldas, E. D., & Recena, M. C P. (2005). Uso de agrotóxicos e suicídios no estado do Mato Grosso do Sul, Brasil. *Cadernos de Saúde Pública*, 21(2), 598-605.

- Prieto, D., & Tavares, M. (2005). Fatores de risco para suicídio e tentativa de suicídio: incidência, eventos estressores e transtornos mentais. *Jornal Brasileiro de Psiquiatria*, 54(2), 146-154.
- Range, L. M., & Knott, E. C. (1997). Twenty suicide assessment instruments: evaluation and recommendations. *Death Studies*, 21, 25-58.
- Rapeli, C. B., & Botega, N. J. (1998). Tentativas de suicídio envolvendo risco de vida: internações em um hospital geral. *Jornal Brasileiro de Psiquiatria*, 47(4), 157-162.
- Teixeira, A. M. F., & Luis, M. A. V. (1997). Distúrbios psiquiátricos, tentativas de suicídio, lesões e envenenamento em adolescentes atendidos em uma unidade de emergência, Ribeirão Preto, São Paulo, 1988-1993. *Cadernos de Saúde Pública*, 13(3), 517-525.
- Urbina, S. (2004). *Essentials of Psychological Testing*. New Jersey: John Wiley & Sons, Inc.
- Werlang, B. S. G., Borges, V. R., & Fensterseifer, L. (2005). Fatores de risco ou proteção para a presença de ideação suicida na adolescência. *Revista Interamericana de Psicología*, 39(2), 259-266.
- Werneck, G. L., Hasselmann, M. H., Phebo, L. B., Vireira, D. E., & Gomes, V. L. O. (2006). Tentativas de suicídio em um hospital geral no Rio de Janeiro, Brasil. *Cadernos de Saúde Pública*, 22(10), 2201-2206.
- World Health Organization (1999). *Figures and Facts about Suicide*. Geneva: WHO.
- World Health Organization (2000a). Preventing suicide: a resource for general physicians. Em *Mental Behavioral Disorders*. [On-Line]. Recuperado de <http://www.who.org.br>
- World Health Organization (2000b). *Prevention for suicidal behaviors: a task for all*. Em *Mental Behavioral Disorders*. [On-Line]. Recuperado de <http://www.who.org.br>

World Health Organization (2002). Multisite intervention study on suicidal behaviours – SUPRE-MISS: protocol of SUPRE-MISS. Geneva: Department of Mental Health and Substance Dependence. [On-Line]. Recuperado de <http://teach-vip.edc.org/documents/Suicide/SupreProtocol.pdf>.

Tabela 1

Matriz de componentes, comunalidades e alfas de Cronbach para cada fator e escala total

| Escala de Razões para Viver $\alpha=0,861$ | Fator 1 $\alpha=0,876$ | Fator 2 $\alpha=0,842$ | Fator 3 $\alpha=0,687$ | h^2 |
|---|---------------------------|---------------------------|---------------------------|-------|
| RFL5 | | | 0,593 | 0,372 |
| RFL9 | | | 0,611 | 0,394 |
| RFL11 | | | 0,598 | 0,436 |
| RFL13 | 0,521 | | | 0,345 |
| RFL16 | | | 0,635 | 0,432 |
| RFL18 | | 0,568 | | 0,408 |
| RFL19 | 0,533 | | | 0,339 |
| RFL20 | 0,670 | | | 0,612 |
| RFL21 | | | 0,537 | 0,421 |
| RFL22 | 0,627 | | | 0,419 |
| RFL24 | 0,522 | | | 0,524 |
| RFL26 | | 0,637 | | 0,412 |
| RFL31 | | 0,536 | | 0,363 |
| RFL32 | 0,673 | | | 0,527 |
| RFL33 | | 0,712 | | 0,530 |
| RFL35 | 0,646 | | | 0,435 |
| RFL36 | 0,534 | | | 0,387 |
| RFL38 | | 0,702 | | 0,516 |
| RFL39 | 0,632 | | | 0,459 |
| RFL40 | 0,716 | | | 0,538 |
| RFL41 | | 0,714 | | 0,545 |
| RFL42 | 0,648 | | | 0,455 |
| RFL43 | | 0,668 | | 0,499 |
| RFL44 | 0,630 | | | 0,463 |
| RFL45 | 0,627 | | | 0,408 |
| RFL46 | | 0,618 | | 0,400 |
| RFL47 | | 0,583 | | 0,375 |
| RFL48 | | 0,624 | | 0,432 |
| Autovalores | 6,67 | 4,12 | 1,67 | |
| % da variância | 23,81 | 14,70 | 5,95 | |

Conclusões e Considerações Finais

Dentro do contexto principal de Avaliação Psicológica, o presente estudo procurou articular e discutir as qualidades psicométricas que um instrumento deve ostentar para que seu uso seja adequado, a saber, a precisão, validade e também a normatização (Noronha & Alchieri, 2004; Noronha & Vendramini, 2003; Pasquali, 1999; Thomas, 2011; Urbina, 2004). Procurou-se explicar, de maneira geral, as principais considerações a respeito, não somente no que concerne à construção, mas à atualização ou manutenção de instrumentos, uma vez que as evidências de validade são cumulativas e não devem cessar com a publicação do teste (Primi, Muniz, & Nunes, 2009; Urbina, 2004). Além disso, procurou-se também focalizar a importância da adequação de testes estrangeiros ao contexto brasileiro, uma vez que a existência de evidências de qualidades psicométricas adequadas ao formato original de um teste estrangeiro, não garante de que ele apresente as mesmas características em outro país, sendo necessários estudos de adaptação ao novo contexto (AERA, APA, & NCME, 1999; Manzi-Oliveira, Balarini, Marques, & Pasian, 2011).

O presente estudo teve como objetivo principal investigar, em diferentes pesquisas, as qualidades psicométricas da Escala Baptista de Depressão – Versão Adulto (EBADEP-A), tendo como base a Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI). Dentro dos objetivos específicos, procurou-se expor os procedimentos de adaptação e composição fatorial da Escala de Razões para Viver, versão brasileira da *Reasons for Living Scale* (RFL). Embora os procedimentos prático-metodológicos tenham sido realizados com base em um único projeto de pesquisa, a estrutura do presente trabalho foi dividida em quatro artigos, três focados na EBADEP-A, e o último destacando a RFL.

A partir do primeiro artigo, intitulado “Escala Baptista de Depressão – Versão Adulto: evidências de validade e consistência interna em amostra de estudantes”, foi possível investigar diferentes evidências de validade baseada na relação com variáveis externas para a EBADEP-A (AERA, APA & NCME, 1999; Urbina, 2004). Foram verificados os valores das correlações entre os valores totais brutos da EBADEP-A e os valores brutos da *Center for Epidemiological Studies – Depression* (CES-D), além de correlações entre a EBADEP-A e os fatores da CES-D, os quais comprovaram a convergência entre os instrumentos. Além disso, as análises de diferenças de média (ANOVA e teste *t de Student*) entre as classificações categóricas dos instrumentos, a partir do qual os grupos com baixas pontuações médias em um instrumento tiveram pontuações significativamente baixas no outro, também indicando convergência entre os testes.

Outras evidências de validade puderam ser investigadas devido às diferenças nos valores médios obtidos por homens e mulheres nos testes, sendo que estas obtiveram valores estatisticamente maiores que aqueles, acompanhando o resultado de outras pesquisas nacionais e internacionais sobre o tema (Angst, & cols., 2002; Aros & Yoshida, 2009; Baptista, 2004; Carrillo, Rojo, & Staats, 2004; Justo & Calil, 2006; Kessler, 2003; Nolen-Hoeksema, Larson, & Graysib, 1999; Sprock & Yoder, 1997; dentre outros). Por meio das análises de consistência interna, com base na TRI, especificamente, pelo exame da presença de funcionamento diferencial de itens (DIF), foi possível perceber alguns itens com maior aderência pelo grupo de mulheres e outros, por homens, mas que podem ser considerados irrelevantes, pois pelo princípio de equidade pode-se concluir por um equilíbrio nos vieses encontrados (Linacre, 2002; 2010; Rueda, 2007). Assim, concluiu-se não somente por índices satisfatórios de consistência interna para EBADEP-A, mas

também que as diferenças de pontuações médias encontradas anteriormente entre homens e mulheres são reais, e não devidas a possíveis características do teste, com favorecimento de um grupo ou outro na pontuação.

O segundo artigo recebeu o título “Análise Fatorial confirmatória da Escala Baptista de Depressão – Versão Adulto”. O estudo fatorial confirmatório foi realizado por meio da comparação entre duas organizações fatoriais encontradas para a EBADEP-A, a primeira, bifatorial, estabelecida durante os estudos iniciais do instrumento por Souza (2010), a qual os sintomas de humor deprimido, perda ou diminuição de prazer, desesperança, desamparo, autoestima rebaixada, sentimento de incapacidade, inadequação e inutilidade, carência/dependência, negativismo, esquiva de situações sociais, queda de produtividade, pensamento de morte, e falta de perspectiva sobre o futuro foram alocados em um fator denominado “Afetivo-Cognitivo” e o segundo, nomeado “Somático”, agregou os itens relativos à alteração de apetite e/ou peso, insônia, lentidão psicomotora e fadiga /perda de energia, além de dificuldade em concentração e irritabilidade. A segunda estrutura estudada, oferecida por Baptista (2012) indicou a EBADEP-A como unifatorial.

Os valores de assimetria (*Skewness*) e de achatamento (*Kurtosis*), em relação a amostra, indicaram um padrão assimétrico aos itens do instrumento. Ao lado disso, os índices de adequação oferecidos por ambas as estruturas fatoriais oferecidas foram muito próximos, o que não permitiu que somente a Análise Fatorial Confirmatória fosse suficiente para concluir pela estrutura mais parcimoniosa. Assim, outros critérios foram estabelecidos, com base em Hair Jr, Anderson, Tatham e Black (2005), a partir dos quais a composição unifatorial proposta por Baptista (2012) foi concluída como a mais coerente com o construto avaliado, dando à EBADEP-A evidências de validade com base na estrutura

interna (Brown, 2006; Hair Jr, & cols., 2005; Pilati & Laros, 2007; Primi, Muniz, & Nunes, 2009).

Ainda com ênfase na EBADEP-A, o terceiro artigo, “Procedimento de normatização para a Escala Baptista de Depressão – Versão Adulto (EBADEP-A): transferência de normas”, foi apresentado em língua inglesa, devido a pretensão à publicação em um periódico internacional. Tal pesquisa foi centrada em um procedimento de normatização baseado na TRI, partindo de dois princípios básicos, um relativo ao pressuposto que a TRI permite, de forma mais acurada que a TCT, a redução de erros de medida na análise das variáveis latentes e, em segundo lugar, à possibilidade de equalização de duas medidas, de forma a torná-las equiparáveis (Embretson & Reise, 2000; Thomas, 2011).

Uma vez que o objetivo do estudo foi transferir a pontuação de corte da CES-D para a EBADEP-A, no que diz respeito a amostras de estudantes, inicialmente foi observado que ambos os instrumentos se adequavam ao modelo de Rasch, devido aos valores médios dos *infits* e *outfits* encontrados, bem como a distribuição das opções de resposta de ambos os testes (Linacre, 2002; 2010), o que tornou admissível a realização dos procedimentos de calibração e equalização, essenciais para a transferência entre as normas (Bauer & Hussong, 2009; Embretson & Reise, 2000; Smith & cols., 2006; Thomas, 2011; Wyse & Reckase, 2011). Por fim, foi estabelecido o ponto de corte igual a 32 para a EBADEP-A.

O último artigo, “Tradução para o Português, Adaptação e Análise Fatorial da *Reasons for Living Scale*” discutiu primordialmente a importância de se adequar um instrumento internacional quando inserido em outro país, a fim de que ele também receba as qualidades psicométricas básicas, dentro do contexto ao qual ele recebeu tradução (Manzi-Oliveira, & cols., 2011; Urbina, 2004). Especificamente sobre a mensuração da

ideação suicida, construto avaliado pela escala, justificou-se a tradução e adaptação do teste para a língua portuguesa devido à escassez de instrumentos adequados para a avaliação no Brasil, dito que somente um teste possui estudos de adaptação e validação, sendo aprovado ao uso clínico profissional pelo Conselho Federal de Psicologia (CFP, 2011; Cunha & Werlang, 1996; Cunha, 2001).

Após os procedimentos de tradução do teste, foram realizados três conjuntos de análises de componentes principais, de forma a explorar suas possíveis estruturas fatoriais. A Escala de Razões para Viver ficou constituída por 28 itens, com base em uma análise com rotação *Varimax* e supressão de cargas fatoriais menores que 0,50, on the one hand isto é, a participação de ao menos 25% de cada item no fator, assim como realizado no estudo de Linehan, Goodstein, Nielsen e Chiles (1983). Foram gerados três fatores, nomeados “Crenças de sobrevivência e enfrentamento”, “Medo da morte e de desaprovação social” e “Motivação de controle externo”. Os índices de confiabilidade oferecidos tanto pela escala total quanto por cada um dos fatores, de acordo com o alfa de Cronbach, foram considerados satisfatórios.

Diante das pesquisas expostas, pode-se perceber que de forma geral, o presente estudo ofereceu as qualidades psicométricas básicas não só à Escala Baptista de Depressão – Versão Adulto, como também à Escala de Razões para Viver, cumprindo os objetivos propostos inicialmente. Partindo do princípio que os procedimentos validação de um instrumento são contínuos, não direcionados somente aos procedimentos de construção do mesmo, os três primeiros estudos puderam oferecer diferentes evidências de validade e consistência interna para a EBADEP-A, complementando o que já foi previamente estabelecido e publicado em seu manual técnico (Baptista, 2012). Além disso, a partir do

contexto de atualização de um teste psicológico, dentro do qual estão inseridos os procedimentos de tradução de um instrumento, o último artigo pode também apresentar uma escala de avaliação da ideação suicida, com a investigação de suas qualidades psicométricas iniciais.

Por fim, devem ser ressaltadas algumas restrições ao presente estudo. De maneira geral, pode-se observar uma limitação comum a quase todos os artigos, qual seja, o uso de uma amostra não clínica, o que pode ter resultado em um “efeito-chão” à EBADEP-A, ou seja, pontuações majoritariamente baixas e por conseguinte, a classificação da maioria dos participantes com ausência de sintomatologia depressiva ou sintomatologia leve.

Além disso, pode ter sido também fator limitador ao estudo a utilização da CES-D como único instrumento de critério. Primeiramente, ela possui ênfase no humor depressivo, enquanto a EBADEP-A se estrutura de forma mais equilibrada de distribuição nos tipos de sintomas. Em segundo lugar, a CES-D, talvez devido ao seu caráter de rastreamento, não possui classificação de gravidade dos sintomas depressivos, mas somente um ponto de corte, o qual indica a presença ou ausência de sintomatologia depressiva, o que pode trazer restrições às interpretações.

Assim, pesquisas futuras com a utilização de populações clínicas, ou seja, com diagnóstico comprovado de depressão, podem ser úteis tanto no que concerne à possibilidade de outras evidências de validade, tipo critério, para a EBADEP-A, quanto a transferência de normas com base em uma população depressiva, gerando novos pontos de corte à escala. Além disso, um diferente grupo amostral pode ser importante na realização de novas análises fatoriais de caráter confirmatório, de forma que a estrutura indicada possa ser corroborada ou confrontada.

Quanto ao estudo sobre o instrumento de ideação suicida, a estrutura de três fatores proposta da RFL, embora teórica e semanticamente mais ajustada, apresentou algumas limitações estatísticas quando comparadas à estrutura de seis fatores, baseado no estudo original de Linehan e colaboradores (1983). Resumidamente, a primeira estrutura demonstrou menor porcentagem de variância total explicada e maior porcentagem de resíduos não redundantes, enquanto a segunda não tenha recebido respaldo teórico e semântico para a divisão fatorial oferecida.

Aconselha-se, portanto, que estudos futuros sejam realizados de forma a dar continuidade à adaptação da Escala de Razões para Viver, aplicando-a juntamente com outros instrumentos, de forma a se buscar diferentes evidências de validade para o instrumento. Complementarmente, uma vez que quase metade dos itens originais foi suprimida da estrutura final, sugere-se também a construção de um instrumento de ideação suicida para o contexto brasileiro.

Referências

- American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME]. (1999). *Standards for Psychology and Educational Testing*. Washington, DC: American Psychology Association.
- Angst, J., Gamma, A., Gastpar, M., Lépine, J-P., Mendlewicz, J., & Tylee, A. (2002). Gender differences in depression. *European Archives of Psychiatry and Clinical Neuroscience*, 252, 201-209.
- Aros, M. S., & Yoshida, E. M. P. (2009). Estudos da depressão: instrumentos de avaliação e gênero. *Boletim de Psicologia*, 59(130), 61-76.
- Baptista, M. N. (2004). *Suicídio e Depressão: Atualizações*. São Paulo: Guanabara Koogan.
- Baptista, M. N. (2012). *Manual Técnico da Escala Baptista de Depressão – Versão Adulto (EBADEP-A)*. São Paulo: Vetor.
- Bauer, D. J., & Hussong, A. M. (2009). Psychometric Approaches for Developing Commensurate Measures Across Independent Studies: Traditional and New Models. *Psychological Methods*, 14 (2), 101-125.
- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.
- Carrillo, J. M., Rojo, N., & Staats, A. W. (2004). Women and vulnerability to depression: some personality and clinical factors. *The Spanish Journal of Psychology*, 7(1), 29-39.
- Conselho Federal de Psicologia (CFP). (2011). Sistema de Avaliação de Testes Psicológicos (SATEPSI). Disponível em: <http://www.pol.org.br/satepsi/>

- Cunha, J. A. (2001). *Manual da versão em português das Escalas de Beck*. Casa do Psicólogo, São Paulo.
- Cunha, J. A., & Werlang, B. C. (1996). Um estudo com a Escala de Desesperança de Beck em grupos clínicos e não-clínicos. *Psico* 27, 189-197.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum.
- Hair Jr., J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise Multivariada de Dados*. Porto Alegre: Bookman.
- Justo, L. P., & Calil, H. M. (2006). Depressão – o mesmo acometimento para homens e mulheres? *Revista de Psiquiatria Clínica*, 33(2), 74-79.
- Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders*, 74, 5-13.
- Linacre J. M. (2002). What do Infit and Outfit, Mean-Squared and Standardized mean? *Rasch Measurement Transactions*, 16 (2), 878. Recuperado de <http://209.238.26.90/rmt/rmt82a.htm>
- Linacre, J.M. (2010). Winsteps® (Version 3.70.0.2) [Computer Software]. Beaverton, Oregon: Winsteps.com.
- Linehan, M. M., Goodstein, J. L., Nielsen, S. L., & Chiles, J. A. (1983). Reasons for Staying Alive. When you are thinking of killing yourself. The Reasons for Living Inventory. *Journal of Consulting and Clinical Psychology*, 51(2), 276-286.
- Manzi-Oliveira, A. B., Balarini, F. B., Marques, L. A. S., & Paisan, S. R. (2011). Adaptação cultural de instrumentos de Avaliação Psicológica: levantamento de estudos realizados no Brasil de 2000 a 2010. *Psico-USF*, 16(3), 367-381.

- Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the Gender Difference in Depressive Symptoms. *Journal of Personality and Social Psychology*, 77(5), 1061-1072.
- Noronha, A. P. P., & Alchieri, J. C. (2004). Conhecimento em Avaliação Psicológica. *Estudos de Psicologia*, 21(1), 43-52.
- Noronha, A. P. P., & Vendramini, C. M. M. (2003). Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica*, 16(1), 177-182.
- Pasquali, L. (1999). *Instrumentos psicológicos: manual prático de elaboração*. Brasília: LabPAM & IBAPP.
- Pilati, R., & Laros, J. A. (2007). Modelos de Equações Estruturais em Psicologia: conceitos e aplicações. *Psicologia: Teoria e Pesquisa*, 23(2), 205-216.
- Primi, R., Muniz, M., & Nunes, C. H. S. S. (2009). Definições contemporâneas de Validade de Testes Psicológicos. Em C. S. Hutz (Org). *Avanços e polêmicas em avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Rueda, F. J. M. (2007). O funcionamento diferencial do item no teste pictórico de memória. *Revista Avaliação Psicológica*, 6(2), 229-237.
- Smith, A. B., Wright, E. P., Rush, R., Stark, D. P., Velikova, G., & Selby, P. J. (2006). Rasch analysis of the dimensional structure of the Hospital Anxiety and Depression Scale. *Psycho-Oncology*, 15, 817-827.
- Souza, M. S. (2010). *Evidências de Validade e Precisão para a Escala de Depressão de Baptista e Sisto (EDEP)*. Tese de doutorado. Programa de Pós-Graduação *Stricto Sensu* em Psicologia, Universidade São Francisco, Itatiba, SP.

- Sprock, J., & Yoder, C. Y. (1997). Women and Depression: an update on the report of the APA Task Force. *Sex Roles, 36*(516), 269-303.
- Thomas, M. L. (2011). The Value of Item Response Theory in Clinical Assessment: A Review. *Assessment, 18*(3), 291-307.
- Urbina, S. (2004). *Essentials of Psychological Testing*. New Jersey: John Wiley & Sons, Inc.
- Wyse, A. E., & Reckase, M. D. (2011). A Graphical Approach to Evaluating Equating Using Test Characteristic Curves. *Applied Psychological Measurement, 35* (3), 217-234.

ANEXOS

Anexo A – Termo de Consentimento Livre e Esclarecido

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (1ª VIA)
ÁREA DE CIÊNCIAS HUMANAS E EXATAS
COMITÊ DE ÉTICA EM PESQUISA – UNIVERSIDADE SÃO FRANCISCO

Pesquisa
Estudos Psicométricos da Escala Baptista de Depressão em Adultos (EBADEP-A)
Responsáveis: Juliana Oliveira Gomes; Makilim Nunes Baptista

Eu _____, portador da carteira de identidade _____, residente na cidade de _____, UF _____, declara ser de livre e espontânea vontade que está participando como voluntário do projeto de pesquisa supra-citado, de responsabilidade do Prof Dr Makilim Nunes Baptista e da doutoranda e pesquisadora Juliana Oliveira Gomes.

O participante declara também estar ciente que:

- 1- O objetivo da pesquisa é obter evidências de validade para a Escala Baptista de Depressão em Adultos (EBADEP-A).
- 2- Durante o estudo os participantes da pesquisa serão solicitados a responder aos seguintes instrumentos: um questionário de identificação, contendo informações gerais como sexo, idade, escolaridade e outras informações relevantes, mantendo, porém o sigilo na pesquisa, a Escala Baptista de Depressão em Adultos (EBADEP-A) e a Escala de Depressão CES-D, que são duas escalas que questionam sobre a presença de sintomatologia depressiva nas últimas semanas e o Inventário de Razões para Viver (RFL) o qual apresenta frases contendo motivos para uma pessoa viver. O tempo estimado para responder os instrumentos é de aproximadamente 30 minutos.
- 3- Obtive todas as informações necessárias para poder decidir conscientemente sobre a minha participação na referida pesquisa.
- 4- A resposta a este(s) instrumento(s)/procedimento(s) não causam riscos conhecidos à minha saúde física e mental, mas podem causar desconforto emocional.
- 5- Estou livre para interromper a qualquer momento minha participação na pesquisa, o que não me causará nenhum prejuízo.
- 6- Meus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos através da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada.
- 7- Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa pelos telefones: (11) 24548981 ou 11 24548028 em Bragança Paulista.
- 8- Poderei entrar em contato com o responsável pelo estudo, Ms Juliana Oliveira Gomes, sempre que julgar necessário pelo telefone (11) 8353-4100.
- 9- Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

Data: ____/____/____

Assinatura do participante:.....

Anexo B – Parecer emitido pelo Comitê de Ética em Pesquisa

Comitê de Ética em Pesquisa – CEP



Bragança Paulista, 07 de Dezembro de 2011

COMITÊ DE ÉTICA EM PESQUISA

Projeto de Pesquisa: ESTUDOS PSICOMÉTRICOS DA ESCALA BAPTISTA DE DEPRESSÃO EM ADULTOS (EBADEP-A).

Curso: Psicologia

Grupo: III

Autor (es): Prof. Makilim Nunes Baptista; Juliana Oliveira Gomes

Instituição: Universidade São Francisco

Protocolo CAAE: 0422.0.142.000-11

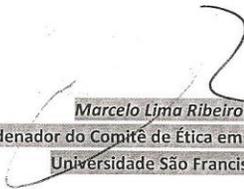
Prezado (a) (s) Pesquisador (a) (s),

O Comitê de Ética em Pesquisa – CEP, da Universidade São Francisco, analisou expressamente dia 07/12/2011 as pendências do projeto de pesquisa supracitado, sob a responsabilidade de Vossa Senhoria.

Este Comitê, acatando o parecer do relator indicado, apresenta-lhe o seguinte resultado:

Parecer: APROVADO

Atenciosamente,


Marcelo Lima Ribeiro
 Coordenador do Comitê de Ética em Pesquisa - USF
 Universidade São Francisco

Qualquer alteração no protocolo de pesquisa, deve ser comunicado e enviado ao CEP-USF.

Ao término do desenvolvimento do estudo o (a) pesquisador (a) responsável deverá enviar ao CEP-USF o relatório consolidado de conclusão do mesmo.

CÂMPUS DE BRAGANÇA PAULISTA Av. São Francisco de Assis, 218 - CEP 12916-900 Fone (11) 4034-8000 - FAX (11) 4034-1825

CÂMPUS DE CAMPINAS Rua Waldemar César da Silveira, 105 - Cura D'Arç CEP 13045-270 (19) 3779-3300

CÂMPUS DE ITATIBA Rua Alexandre Rodrigues Barbosa, 45 - CEP 13251-900 Fone (11) 4534-8000 - FAX (11) 4524-1933

CÂMPUS DO PARI - SÃO PAULO Rua Hannemann, 352 - Pari - CEP 03031-040 Fone (11) 3315-2000 - FAX (11) 3315-2036

Anexo C – Questionário de Identificação

Data: ____/____/____

1 - Sexo: ¹() Masculino ²() Feminino 2- Idade: 3- Curso:

4 - Estado Civil:

¹() Solteiro ⁴() Desquitado/ Divorciado
²() Casado ⁵() Recasado
³() Viúvo ⁶() Outros:

5 – Indique o nível de escolaridade do chefe da família:

¹() Analfabeto/ fundamental incompleto ⁴() Médio completo / Superior incompleto
²() Fundamental I (1ª a 4ª) completo/ ⁵() Superior completo
Fundamental II (5ª a 8ª) incompleto
³() Fundamental II completo/
Médio incompleto

6 – Se você não for o chefe da família, indique seu grau de escolaridade:

¹() Analfabeto/ fundamental incompleto ⁴() Médio completo / Superior incompleto
²() Fundamental I (1ª a 4ª) completo/ ⁵() Superior completo
Fundamental II (5ª a 8ª) incompleto
³() Fundamental II completo/
Médio incompleto

7 – Assinale com um X a quantidade dos itens que você possui em casa:

| Itens | Não tem | 1 | 2 | 3 | 4 ou + |
|--|---------|---|---|---|--------|
| Televisão | | | | | |
| Rádio | | | | | |
| Banheiro | | | | | |
| Automóvel | | | | | |
| Empregada mensalista | | | | | |
| Aspirador de pó | | | | | |
| Maquina de lavar | | | | | |
| Videocassete e/ou DVD | | | | | |
| Geladeira | | | | | |
| Freezer (aparelho independente ou parte da geladeira duplex) | | | | | |

8- Trabalha atualmente?

1() Não 2() Sim. Qual profissão? : _____

9 – Alguém da sua família já foi diagnosticado com depressão, por psicólogo ou psiquiatra?

1() Não 2() Sim. Quem? _____

10 – Você já passou por avaliação psiquiátrica ou psicológica?

1() Não 2() Sim, com diagnóstico de depressão
3() Sim, sem nenhum diagnóstico 4() Sim, com outro diagnóstico. Qual? _____

11 –Atualmente está passando por tratamento para depressão?

psicológico 1() Não 2() Sim Há quanto tempo? _____
medicamentoso 1() Não 2() Sim Há quanto tempo? _____

12- Se você está diagnosticado com algum transtorno depressivo, acredita que atualmente os sintomas depressivos se encontram em qual nível?

¹Leve () ²() Moderado ³() Severo

Anexo D – Escala de Razões para Viver

UNIVERSIDADE DE WASHINGTON
CLINICA DE PESQUISA E TERAPIA COMPORTAMENTAL
Linehan et. al., 1983.

UNIVERSIDADE SÃO FRANCISCO
PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM PSICOLOGIA
Baptista & Gomes, 2013

Escala de razões para viver

INSTRUÇÕES: Muitas pessoas já pensaram em suicídio pelo menos uma vez. Outros nunca o consideraram. Se você pensou nisso ou não, nós estamos interessados nas razões que você teria para **não** cometer suicídio se o pensamento lhe ocorreu ou se lhe foi sugerido por alguém.

Nas páginas seguintes há razões dadas por pessoas do porque não cometer o suicídio. Nós gostaríamos de saber o quão importante cada uma dessas possíveis razões seriam para você, a essa altura de sua vida, como uma razão para não se matar. Por favor, dê sua avaliação no espaço a esquerda de cada questão.

Cada razão pode ser avaliada de 1 (Não tão importante) a 6 (Extremamente Importante). Se uma razão não se aplica a você ou se você não acredita que o tópico seja verdadeiro, então ele não é assim tão importante você deve colocar (1). Por favor use toda a gama de escolhas e não somente as opções do meio (2,3,4,5) ou somente as extremidades (1,6).

Em cada espaço, coloque o número para indicar a importância para você de cada razão para não se matar.

1. Sem importância (como uma razão para **não** me matar, ou isso não se aplica a mim, eu não acredito em tudo isso).
2. Não tão importante.
3. Algo sem Importância.
4. Pouco importante.
5. Muito importante.
6. Extremamente Importante (Como uma razão para **não** me matar, eu creio muito nisso, e isso é muito importante).

Mesmo que você nunca tenha, ou realmente acredite que você nunca considerou seriamente se matar, é realmente muito importante que você avalie cada razão. Nesse caso, considerando a base de **porque se matar não é ou não seria nunca uma alternativa para você.**

Em cada espaço coloque um número para indicar a importância para você de não se matar.

- | | |
|---------------------------------|-----------------------------|
| 1. Sem importância. | 4. Pouco importante. |
| 2. Não tão importante. | 5. Muito importante. |
| 3. Algo sem Importância. | 6. Extremamente Importante. |
-

- _____ 1. Eu tenho uma responsabilidade e compromisso com a minha família.
- _____ 2. Eu acredito que posso aprender a ajustar ou lidar com meus problemas.
- _____ 3. Eu acredito que tenho controle sobre minha vida e destino.
- _____ 4. Eu tenho o desejo de viver.
- _____ 5. Eu acredito que apenas Deus tem o poder de encerrar uma vida.
- _____ 6. Eu tenho medo da morte.
- _____ 7. Minha família pode acreditar que eu não os amei.
- _____ 8. Eu não acredito que as coisas se tornam tão miseráveis e sem esperança suficiente para que eu preferisse estar morto.

- | | |
|---------------------------------|-----------------------------|
| 1. Sem importância. | 4. Pouco importante. |
| 2. Não tão importante. | 5. Muito importante. |
| 3. Algo sem importância. | 6. Extremamente Importante. |
-

- _____ 9. Minha família depende de mim e precisa de mim.
- _____ 10. Eu não quero morrer.
- _____ 11. Eu quero ver o crescimento de meus filhos.
- _____ 12. A vida é tudo o que temos e é melhor do que nada.
- _____ 13. Eu tenho planos futuros, estou almejando a realização dos mesmos.
- _____ 14. Não importa o quão ruim eu me sinta, eu sei que isso não vai durar.
- _____ 15. Tenho medo do desconhecido.
- _____ 16. Eu amo e curto tanto minha família, que eu não poderia viver sem eles.
- _____ 17. Eu quero experimentar tudo o que a vida tem para me oferecer e há muitas experiências que eu ainda não tive e que eu quero ter.
- _____ 18. Tenho medo que meu método de me matar possa falhar.
- _____ 19. Eu me preocupo o suficiente sobre mim mesmo para viver.
- _____ 20. A vida é tão bela e preciosa para acabar com ela.
- _____ 21. Não seria justo deixar as crianças para outros tomarem conta.
- _____ 22. Eu acredito que eu posso achar outras soluções para meus problemas.
- _____ 23. Tenho medo de ir para o inferno.
- _____ 24. Eu tenho amor a vida.
- _____ 25. Eu sou muito estável para me matar.
- _____ 26. Eu sou um covarde e não tenho coragem de fazê-lo.
- _____ 27. Minhas crenças religiosas proíbem isso.
- _____ 28. O efeito sobre os meus filhos pode ser prejudicial.
- _____ 29. Eu sou curioso em saber o que vai acontecer no futuro.
- _____ 30. Eu vou machucar muito minha família e eu não quero que eles sofram.
- _____ 31. Estou preocupado com o que os outros pensariam de mim.
- _____ 32. Eu acredito que tudo tem um jeito de trabalhar para o melhor.
- _____ 33. Eu não conseguiria decidir como, quando e onde fazê-lo.
- _____ 34. Eu considero isso moralmente incorreto.
- _____ 35. Eu ainda tenho muitas coisas a fazer.
- _____ 36. Eu tenho coragem de encarar a vida.
- _____ 37. Estou contente e satisfeito com minha vida.
- _____ 38. Tenho medo do "ato" real de me matar (a dor, sangue, violência).
- _____ 39. Eu acredito que me matar não seria realmente realizar ou resolver nada.
- _____ 40. Eu tenho esperança que as coisas vão mudar e o futuro será mais feliz.
- _____ 41. Outras pessoas pensariam que eu sou fraco e egoísta.
- _____ 42. Eu tenho um impulso interior para sobreviver.
- _____ 43. Eu não gostaria que as pessoas pensassem que eu não tive controle sobre minha vida.
- _____ 44. Eu acredito que eu possa encontrar um propósito na vida, uma razão para viver.
- _____ 45. Não vejo nenhuma razão para apressar a morte.
- _____ 46. Eu sou tão incapaz que o meu método não funcionaria.
- _____ 47. Eu não gostaria que minha família se sentisse culpada depois.
- _____ 48. Eu não gostaria que minha família pensasse que eu fui covarde ou um egoísta.