UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE PSICOLOGIA

SONIA MARIA DA SILVA

ESCALA DE AVALIAÇÃO DE RESULTADO (OUTCOME QUESTIONNAIRE) - OQ 45.2: VALIDADE E PRECISÃO

(VERSÃO CORRIGIDA)

São Paulo

2013

SONIA MARIA DA SILVA

ESCALA DE AVALIAÇÃO DE RESULTADOS (OUTCOME QUESTIONNAIRE) - OQ 45.2: VALIDADE E PRECISÃO

(Versão corrigida)

Tese apresentada ao Instituto de Psicologia da Universidade de São Paulo para obtenção do título de Doutora em Psicologia.

Área de concentração: Programa de Pós-Graduação em Psicologia Escolar e Desenvolvimento Humano.

Orientadora: Profª Drª Irai Cristina Boccato Alves

São Paulo

2013

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Catalogação na publicação

Biblioteca Dante Moreira Leite

Instituto de Psicologia da Universidade de São Paulo

Silva, Sonia Maria da.

Escala de Avaliação de Resultados (Outcome Questionnaire) - OQ-45.2: Validade e precisão / Sonia Maria Da Silva; orientadora Irai Cristina Boccato Alves. -- São Paulo, 2013.

98 f.

Tese (Doutorado – Programa de Pós-Graduação em Psicologia. Área de Concentração: Psicologia Escolar e do Desenvolvimento Humano) – Instituto de Psicologia da Universidade de São Paulo.

1. Avaliação psicológica 2. Avaliação de processos e resultados 3. Validade do teste 4. Precisão do teste I. Título.

BF176

FOLHA DE APROVAÇÃO

SILVA, S.M. (2013). **Escala de Avaliação de Resultados – (Outcome Questionnaire) - OQ-45.2: Validade e Precisão.**

Tese apresentada ao Instituto de Psicologia da Universidade de São Paulo para obtenção do título de Doutora em Psicologia.

Área de concentração: Psicologia Escolar e do Desenvolvimento Humano.

Aprovada em: ___/___/____

Banca Examinadora:

Profª. Dra Patrícia Waltz Schelini

Instituição: Universidade F. São Carlos. Assinatura_____

Profª. Drª.Elisa Pizão Médici Yoshida

Instituição: PUC CAMPINAS. Assinatura_____

Profª. Drª Audrey Setton Lopes de Souza

Instituição: Universidade de São Paulo. Assinatura_____

Profª. Drª Sonia Beatriz Meyer

Instituição: Universidade de São Paulo. Assinatura_____

Prof. Dra. Irai Cristina Boccato Alves

Instituição: Universidade de São Paulo. Assinatura_____

A minha neta Sophia Flora que me faz redescobrir a “velha” infância.

Aos meus filhos Angélica, Gabriel e Apolo.

A minha Felismina e meu pai José (in memorian).

Ao Paulo, com amor.

AGRADECIMENTOS

À Profa. Dra. Irai Cristina Boccato Alves, “orientadora” no sentido literal da palavra, pela competência, dedicação, carinho, paciência e disponibilidade, sem as quais não teria sido possível a realização e finalização deste trabalho, vou ser sempre grata pela sua generosidade em compartilhar conhecimentos, por sua sensibilidade e voto de confiança.

À Universidade Guarulhos pelo apoio em permitir a realização desta pesquisa, autorizando a coleta de dados.

Aos amigos da Universidade Guarulhos que me perdoem não identificá-los, mas o tempo não permitiu. Vocês fazem parte da minha história profissional e pessoal.

Aos professores da Universidade Presbiteriana Mackenzie com quem compartilho muitas horas do meu cotidiano, os meus agradecimentos (como vocês bem sabem esta é a ultima coisa que escrevemos, não nomeá-los foi uma estratégia para não correr o risco de esquecer amigos que são tão importantes para mim).

Aos meus alunos que diretamente e indiretamente contribuíram para a realização deste trabalho.

À minha família, por estar ausente nos últimos meses, nas festas de aniversário, Natal e Ano Novo. Em especial a minha mãe, com toda a sua sabedoria, me dá forças.

Aos meus irmãos Vera, Ana, Arnaldo, João, Eduardo e Marco Paulo.

"Aprender é a única coisa de que a mente nunca se cansa, nunca tem medo e nunca se arrepende."

Leonardo da Vinci

RESUMO

SILVA, Sonia Maria. (2013). **Escala de Avaliação de Resultados (Outcome Questionnaire) - OQ-45.2: Validade e Precisão**. Tese (Doutorado) - Universidade de São Paulo, Programa de Pós Graduação Stricto Sensu em Psicologia Escolar e Desenvolvimento Humano, São Paulo. 98 p.

As pesquisas para padronização de instrumentos de avaliação de resultados terapêuticos têm sido produtivas, uma vez que a aproximação entre a pesquisa e a prática clínica permite ao profissional perceber a importância dessa avaliação de processo e de resultados. Com isto, a necessidade dos psicólogos contarem com instrumentos de avaliação, capazes de auxiliá-los nessa tarefa, tem se mostrado cada vez mais essencial. No entanto, a carência de instrumentos válidos para a realidade brasileira impõe limitações a estes profissionais. O Outcome Questionnaire - OQ-45.2 é um dos instrumentos utilizados para avaliar os ganhos obtidos pelo paciente na psicoterapia, com pesquisas realizadas em vários países. Trata-se de uma escala do tipo *Likert* de cinco pontos, dividida em três subescalas: desconforto subjetivo (SD), relações interpessoais (IR) e desempenho do papel social (SR). Esta pesquisa teve o objetivo de avaliar as propriedades psicométricas da versão em português do OQ-45.2. A amostra foi composta por 419 participantes adultos, subdivididos em dois grupos, um de pacientes (N = 59) e outro de não pacientes (N = 360), sendo 156 homens e 263 mulheres, com idades variando de 18 a 78 anos. O estudo de precisão do OQ-45.2 foi realizado pelo método de teste-reteste, com intervalo de 7 a 14 dias, e o alfa de Cronbach. Para a escala total a precisão pelo reteste foi de 0,895 e para as subescalas variou de 0,756 a 0,883, indicando estabilidade temporal satisfatória. O alfa de Cronbach para a escala total foi de 0,95, semelhante ao do estudo americano (0,93). Os dados de validade foram obtidos pela correlação entre o resultado global e os das subescalas, por meio da comparação entre grupos contrastantes de pacientes e não pacientes e da validade simultânea com a Escala de Avaliação de Sintomas (EAS-40), o Inventário de Depressão de Beck (BDI II) e o Questionário Geral de Saúde (QSG). As correlações das subescalas com a pontuação total e entre as subescalas foram significantes, mas a subescala SR mostrou correlações menores que as outras. Foram obtidas diferenças estatisticamente significantes ($p < 0,001$) entre as médias dos dois grupos, o que mostra a sensibilidade da escala para avaliação dos pacientes. O resultado do OQ – 45.2 se correlacionou significativamente com os três instrumentos utilizados para o estudo de validade, sendo que as correlações com os escores totais dos instrumentos foram: 0,80 com a EAS-40, 0,83 com o BDI-II e 0,88 com o QSG. Também foram calculadas as correlações entre cada item com o escore total do OQ-45.2, tendo sido encontrados quatro itens com correlações menores do que 0,20. Assim os estudos psicométricos indicaram a adequação da escala para o uso no Brasil, embora ainda seja necessária a confirmação de sua estrutura fatorial e de estudos referentes a mudanças obtidas na psicoterapia.

Palavras-chave: Avaliação psicológica, Avaliação de processos e resultados; Validade do teste; Precisão do teste.

ABSTRACT

SILVA, S. M. (2013). *Outcome Questionnaire (OQ-45.2): Validity and Reliability.* Tese de Doutorado. Instituto de Psicologia. Universidade de São Paulo.

Researches for standardization of therapeutic assessment tools have been productive, since the connection between research and clinical practice allows the professional to realize the importance of the assessment process and results. In this way, the psychologists need of available assessment tools, which can assist them in this task, has become ever more essential. However, the lack of valid instruments to the Brazilian reality imposes limitations to these professionals. The Outcome Questionnaire - OQ-45.2 is one of the instruments used to measure the patient's progress in psychotherapy and it was studied in several countries. It is a Likert five points scale, divided into three subscales: Symptom Distress (SD), Interpersonal Relationships (IR) and Social-Role performance (SR). The goal of this research was to assess the psychometric properties of the OQ-45-2 Portuguese version. The sample was composed by 419 adult participants, divided in two groups, one group of patients (N = 59) and the other of non-patients (N = 360), 156 men and 263 women, age ranging from 18 to 78 years. The reliability study of the OQ-45.2 was conducted by the test-retest method, using intervals ranging from 7 to 14 days, and by the Cronbach's alpha coefficient. The retest reliability for the total score was 0.895 and for the subscales ranged from 0.756 to 0.883, indicating satisfactory temporal stability. The Cronbach's alpha for the total scale was 0.95, similar to the American study (0.93). The validity data were obtained by correlations between the total score and the subscales scores; by comparing contrasting groups of patients and non patients and concurrent validity with the Symptom Assessment Scale (EAS-40) Beck Depression Inventory (BDI II) and the General Health Questionnaire (GHQ). Correlations of the subscales with the total score and subscales scores were significant, but the SR subscale showed lower correlations than others. Differences between the means of two groups were significant statistically ($p < 0.001$), which shows the sensitivity of the scale to assess patients. The results of OQ-45.2 correlated significantly with the three tests used in validity study and correlations between the total scores of the tests were 0.80 with the EAS-40, 0.83 with the BDI-II and 0.88 with the GHQ. Correlations total of total score of OQ-45.2 with each item were also calculated and were found four items correlations lower than 0.20. Therefore, the psychometric studies indicated the appropriateness of the scale for use in Brazil, although it is still necessary the confirmation of its factor structure and studies regarding changes obtained in psychotherapy.

Key words: Psychological Assessment; Processes and Outcomes Assessment; Test Validity; Test Reliability.

SUMÁRIO

1	**INTRODUÇÃO**	14
2	**FUNDAMENTAÇÃO TEÓRICA**	19
	Pesquisa em Psicoterapia	19
	Instrumentos para avaliação da psicoterapia	26
	Características psicométricas dos testes psicológicos	29
	Validade	30
3	**OUTCOME QUESTIONNAIRE – OQ 45.2**	32
	Desconforto Subjetivo (SD)	35
	Relações Interpessoais (IR)	36
	Papel Social (SR)	37
4	**APLICAÇÃO E AVALIAÇÃO DO OQ 45.2**	38
	Padronização Americana	39
	Estudos Internacionais	41
5	**OBJETIVOS**	55
	Objetivo Geral	55
	Objetivos específicos	55
6	**MÉTODO**	56
	Participantes	56
	Instrumentos	61
	Escala de Avaliação de Sintomas (EAS-40)	61
	Inventário de Depressão de Beck (BDI-II)	61
	Questionário Geral de Saúde (QSG)	61
	Procedimentos	62
7	**RESULTADOS E DISCUSSÃO**	66
	Precisão de Teste-Reteste	66
	Validade	68
	Grupos Contrastantes	71
	Validade Concorrente	74
8	**CONSIDERAÇÕES FINAIS**	82
	REFERENCIAS	84
	ANEXOS	96

LISTA DE TABELAS

TABELA 1	Dados demográficos da amostra total (N = 419).	56
TABELA 2	Dados demográficos dos grupos de pacientes e não-pacientes.	57
TABELA 3	Distribuição dos grupos de pacientes e não-pacientes de acordo com a ocupação atual.	59
TABELA 4	Classificação da classe econômica da amostra segundo o Critério de Classificação Econômica Brasil da ABEP.	60
TABELA 5	Total de testes respondidos pela amostra total e por sexo	62
TABELA 6	Comparação da Precisão pelo Reteste das amostras de estudantes universitários (grupo não-clínico).	66
TABELA 7	Coeficiente de Alfa para OQ 45.2 Total e cada subescala nas amostras brasileira e americana de estudantes universitários e pacientes.	67
TABELA 8	Coeficiente Alfa de Cronbach para OQ 45.2 e para cada subescala das amostras brasileiras e americana de pacientes.	68
TABELA 9	Correlação entre pontuação total e as subescalas na amostra total (N = 419) e não-pacientes.	69
TABELA 10	Correlações entre as subescalas SD, IR, SR e escore total para cada grupo.	70
TABELA 11	Médias, desvio padrão e teste *t* entre os grupos de pacientes e não-pacientes.	71
TABELA 12	Diferenças entre gêneros nas Médias e Desvio Padrão para o OQ Total	72
TABELA 13	Correlação entre os escores total e as dimensões da EAS-40 com escore total e dimensões do OQ 45.2 (NP = 322, P = 51).	73
TABELA 14	Correlações entre a EAS-40 e o OQ 45.2 para os grupos de pacientes (N = 51) e não pacientes (N = 322).	74
TABELA 15	Correlações entre BDI –II e o OQ 45.2 (N = 189).	76
TABELA 16	Correlações entre BDI-II e o OQ 45.2 dos dois grupos (P = 22 e NP = 167).	77

TABELA 17 Correlações entre QSG e OQ 45.2 (N = 115). 78

TABELA 18 Correlações entre QSG e OQ 45.2 dos grupos de não-pacientes (N = 95) e pacientes (N = 20). 79

LISTAS DE ANEXOS

ANEXO 1	Termo de Consentimento Livre e Esclarecido: Pacientes	96
ANEXO 2	Termo de Consentimento Livre e Esclarecido: Não pacientes	97
ANEXO 3	Aprovação da Pesquisa: Comitê de Ética	98

INTRODUÇÃO

Nos últimos anos observa-se um avanço na área de avaliação psicológica, pois cada vez mais professores e pesquisadores se mostram preocupados com o ensino, e com treinamento na aplicação, avaliação, interpretação dos instrumentos de avaliação psicológica e atualização dos estudos. As propriedades psicométricas tais como validade e precisão passam a ser tema de discussões, especialmente com a publicação de Resolução do Conselho Federal de Psicologia (CFP) Nº 002/2003, que considera a necessidade de aprimorar os instrumentos e procedimentos técnicos do trabalho dos psicólogos bem como de fazer a revisão periódica dos mesmos.

As determinações da Resolução CFP N.º 002/2003 têm como objetivo garantir serviços com qualidade técnica e ética à população usuária dos serviços de Psicologia (Conselho Federal de Psicologia, 2003). Esta Resolução estabelece os requisitos mínimos e obrigatórios para todos os instrumentos de avaliação psicológica, inclusive os que utilizam questões de múltipla escolha e outros similares, tais como "acerto e erro", "inventários" e "escalas". Também estão sujeitos aos requisitos estabelecidos nesta Resolução os testes estrangeiros de qualquer natureza, traduzidos para o português, que devem ser padronizados a partir de estudos realizados com amostras brasileiras, apresentar evidências empíricas de validade, precisão e dados normativos para o ambiente cultural onde serão utilizados, relatar de modo claro e exaustivo as características da amostra de padronização, possibilitando a avaliação do nível de representatividade do grupo de referência.

Dentro da gama de instrumentos de medida usados em avaliação psicológica, aqueles chamados de autorrelato têm recebido destaque, pois por serem respondidos pelo próprio indivíduo, permitem aferir a maneira pela qual ele relata a percepção sobre o próprio estado. Esses instrumentos vêm sendo cada vez mais usados na área de avaliação de processo e resultados de psicoterapia, uma vez que essas medidas parecem contribuir de modo mais significativo para a avaliação das mudanças do paciente, (Harmon, Hawkins, Lambert, Slade & Whipple, 2005). São escalas simples e que acarretam menos esforços, dado que conseguem aproximar a pesquisa da prática clínica (Lambert, Hansen & Finch, 2001).

Apesar das críticas feitas às escalas de autorrelato quanto as suas limitações, respostas tendenciosas dadas na direção do comportamento esperado socialmente, como são instrumentos de fácil entendimento para o examinando, que fornece informações sobre si mesmo, evita julgamentos subjetivos do avaliador, bem como trazem resultados que permitem a realização de pesquisas quantitativas com maior número de participantes, de

modo que esses resultados possam ser generalizados.

A percepção do paciente sobre o seu estado emocional complementa a avaliação do psicólogo durante o psicodiagnóstico e integra uma visão mais ampla sobre seu estado psicológico. Uma percepção multilateral do conflito psíquico é defendida há bastante tempo na literatura internacional relativa à pesquisa em psicoterapia, que se dedica à investigação das mudanças psíquicas. Apesar da extensa tradição na pesquisa dos fatores que podem contribuir para ganhos na psicoterapia, os estudos continuam, pois ainda há questões não respondidas integralmente (Strupp, Horowitz & Lambert, 1997). Segundo Markin, Mccarthy e Barber (2013) sabe-se relativamente pouco sobre como os fatores da relação terapêutica se inter-relacionam e como, juntos, eles impactam o processo e o resultado da psicoterapia.

No Brasil observa-se o aumento de pesquisas relacionadas aos resultados do tratamento psicológico, entretanto, segundo Serralta, Nunes e Eizirik (2007), tem havido pouca sistematização quanto à utilização de instrumentos para avaliação de processos psicoterapêuticos, assim como no tocante às formas de atendimento e aos resultados de psicoterapias. Os mesmos autores acrescentam que a pesquisa em psicoterapia se encontra em estado inicial e ainda faltam estudos sistemáticos, em profundidade sobre esse tema.

A delimitação de um tema no campo da pesquisa em psicoterapia não é uma tarefa fácil, uma vez que o desdobramento do mesmo em partes requer do pesquisador conhecimento e experiência prática em psicoterapia. Segundo Meyer, Oshiro, Donadone, Mayer e Starling (2008) a complexidade da pesquisa em psicoterapia se relaciona à peculiaridade da interação terapeuta-cliente, que dificilmente se repete em outros atendimentos, assim os resultados de estudos de casos únicos não são suficientes para a generalização para outras amostras.

Para oferecer um panorama sobre as tendências de pesquisas em psicoterapia em outros países, Enéas (2008) analisou nove seções especiais do *Journal of Consulting and Clinical (JCCP)*, da *American Psychological Association* (APA) referentes ao período de 1981 a 1994. A autora destaca que na literatura há uma preocupação em buscar evidências científicas,

> principalmente para as pesquisas de eficácia que esbarram nas limitações dos delineamentos de pesquisa e nos modelos estatísticos existentes que não se prestam a abranger a complexidade da interação das variáveis do campo (p. 115).

Os pesquisadores buscam identificar quais são os efeitos que a psicoterapia produz.

Que tipo de tratamento psicológico é mais efetivo para determinada psicopatologia? Quais tipos de intervenções são mais adequados para o paciente? Quais são os fatores compartilhados por qualquer modelo de psicoterapia? Em quais momentos evolutivos a psicoterapia é mais adequada? Entre os fatores comuns, quais têm mais influência nos resultados, a aliança terapêutica, modalidades de intervenções ou as capacidades do próprio paciente?

Frente ao montante de perguntas ainda não respondidas constata-se a necessidade de fomentar pesquisas nesta área desde o início da graduação. Os supervisores contribuem para formação dos estagiários de Psicologia, que inicialmente se dá por meio da apresentação do conhecimento consolidado, que leva ao desenvolvimento das habilidades e competências necessárias para o exercício profissional. Também ajudam no desenvolvimento de atitudes que possibilitem uma integração entre a prática clínica e o conhecimento, buscando as melhores estratégias disponíveis para uma atuação segura e ética nas ações. Aliar a pesquisa à prática clínica constitui uma tarefa fundamental que precisa ser incentivada ao longo da formação do psicólogo. No âmbito da clínica-escola é auxiliar no aprimoramento de principais funções, tanto na capacitação do aluno para a atuação profissional, como no desenvolvimento de seu raciocínio clínico, aproximando a pesquisa da prática clínica, constituindo um espaço de produção de conhecimento e de prestação de serviço psicológico a diversas camadas da população.

A utilização de instrumentos para avaliar o progresso do paciente em diversos momentos da psicoterapia complementa a formação do profissional. Neste sentido, a clínica–escola é um espaço reservado às práticas psicológicas, representa a porta de entrada do aluno de Psicologia para o "mundo profissional" (Yoshida, 2005) e, na formação de psicólogos, é fundamental desenvolver uma prática associada ao uso do conhecimento publicado em periódicos da área. Yoshida (2005) adverte que, para cumprir esta função de formação acadêmica, a clínica-escola deve se constituir também em um espaço de pesquisa do psicólogo, uma vez que oferece uma fonte privilegiada de dados que orientam a reflexão sobre os métodos empregados e que podem contribuir para o desenvolvimento de novos conhecimentos.

Outro aspecto importante na formação é o estabelecimento de parcerias de investigação entre o setor acadêmico e os serviços de saúde público e privado. São evidentes os benefícios desta parceria, uma vez que as práticas dos serviços de saúde psicológicos seriam dotadas de fundamentação empírica e as teorias teriam uma validade externa (Machado, Klein, & Farate, 2005). Na atualidade, as clínicas-escola começam a desenvolver

projetos para coleta de dados sistematizada de modo a obter informações úteis para os alunos em formação, buscando referências e *feedback* que possibilitem a melhoria dos tratamentos oferecidos à comunidade.

A proposta deste trabalho se insere em um tema de relevância tanto no campo internacional, em que as pesquisas para padronização de instrumentos de avaliação de resultados terapêuticos têm sido profícuos, como no contexto nacional, em função do número ainda reduzido de investigações na área. Entretanto é fundamental ressaltar o que tem sido produzido até o momento, como os estudos de Azevedo (2006; 2009); Enéas (2003); Freire (2006); Gatti (1994); Honda (2010); Jung (2007); Laloni (2008); Massei (2009); Rocha (2002); Sanches (2009) e Yoshida (2000).

Nos estudos sobre avaliação de psicoterapia em nosso país, segundo Santeiro, (2005) predominam a utilização de escalas clínicas. Foram publicados 14 artigos em dois periódicos nacionais, enquanto que no mesmo período (1980/2002), em periódicos estrangeiros foram encontrados 117 estudos.

Um dos instrumentos mais utilizado em outros países para verificar progressos em psicoterapia é o *Outcome Questionnaire-OQ-45.2* (Lambert et al., 2004). No site *OQ Measures - The Measure of Mental Health Vital Signs* são referenciados 73 artigos relacionados a pesquisas de psicoterapia, tendo como instrumento o *OQ-45.2.* A utilização desta escala é de grande valia para a prática clínica e, em contexto educacional, na formação de psicólogos e sistematização da pesquisa na clínica-escola.

Em outros países como Itália, Alemanha e Holanda, as pesquisas visam a organização de baterias para avaliação de problemas específicos e de resultados de tratamentos para o aperfeiçoamento da psicoterapia (Strupp, Horowitz & Lambert, 1997). Destacam-se os estudos de Lo Coco et al. (2008) e de Jong et al. (2007), nos quais os resultados mostraram que as propriedades psicométricas das versões holandesa e italiana são adequadas e similares às do estudo original do instrumento.

Em pesquisas realizadas no Chile, os resultados *mostraram que o instrumento OQ-45.2 é válido em sua capacidade de medir resultados terapêuticos* (Rodríguez, 2000, p.103). Além da capacidade de discriminar distintos subgrupos de pesquisa, o instrumento também foi usado para avaliar a qualidade do tratamento de pacientes hospitalizados (Correa, et al., 2006).

A proposta da presente pesquisa foi estimar a precisão e validade da Escala de Avaliação de Resultados de Psicoterapia-*Outcome Questionnaire* (*OQ-45.2*) para a realidade brasileira. Trata-se de um instrumento que não é teoricamente orientado e, assim, pode ser

aplicado a psicoterapias de diferentes fundamentações teóricas, além de apresentar evidências de validade e precisão, em sua versão original.

As motivações para este estudo provêm das experiências como psicóloga e supervisora responsável pela formação de psicólogos na área clínica e, ao mesmo tempo, pelos pacientes que são atendidos pelos estagiários na clínica-escola. Aliar a pesquisa à prática clínica é uma forma de desenvolver no aluno a percepção quanto à importância da avaliação de processo e dos resultados de psicoterapias. A escolha do Outcome Questionnaire se deu em função do número de pesquisas publicadas em diferentes países, tanto estudos avaliando as propriedades psicométricas do instrumento, como pesquisas de avaliação de processos e resultados de psicoterapia. Este trabalho pretende contribuir para oferecer um instrumento que pode facilitar a sistematização de pesquisas em diferentes regiões do país, promovendo maior intercâmbio entre os pesquisadores e integração dos resultados, dentro e fora do país.

FUNDAMENTAÇÃO TEÓRICA

Pesquisa em Psicoterapia

O interesse pela pesquisa em psicoterapia tem crescido em nosso país e, desde que a profissão do psicólogo foi reconhecida no Brasil na década de 60, existe uma preocupação, por parte do Conselho Federal de Psicologia, de delimitar o campo de atuação deste profissional. Por esta razão foi publicada a Resolução 010/2000, que especifica e qualifica a psicoterapia como prática do psicólogo, que propôs:

> Incentivar e promover pesquisas e publicações científicas no campo das psicoterapias; promover maior aproximação da academia com a prática profissional; incentivar a pesquisa e iniciação científica, inclusive no campo da psicoterapia, para desenvolver o olhar de psicólogo e pesquisador (CFP, 2009, p. 12).

No Brasil a pesquisa associada à prática clínica ainda é muito incipiente, o que não ocorre em outros países. Yoshida (2012, p.15) ressalta que "*os pesquisadores de todo o mundo continuam a buscar novas evidências que permitam encontrar os ingredientes ativos em uma psicoterapia"*. Procura-se identificar quais fatores são determinantes nas mudanças obtidas pelo paciente durante o curso do tratamento e em que medida tais mudanças dependem do paciente, do método utilizado na prática da psicoterapia ou do terapeuta. A relevância do papel do psicoterapeuta é expressiva, entretanto a capacidade de empatia, de desenvolver a aliança terapêutica e fazer intervenções ajustadas com momento do tratamento pode aumentar as possibilidades de sucesso da psicoterapia (Yoshida, 2012).

A ideia de que há um distanciamento entre a pesquisa e a psicoterapia depende de conceito de pesquisa. Simon (2010) supõe que a confusão entre pesquisa e tratamento tenha sido feita por Freud, no artigo *Recomendações aos médicos que exercem a Psicanálise*, pois ao mesmo tempo em que "*em sua execução, pesquisa e tratamento coincidem, não obstante, após certo ponto, a técnica exigida por uma opõe-se à requerida pelo outro"* (Freud, 1912/1969, p. 152). O tratamento em si pode ter uma função de instrumento de pesquisa aplicada. "O *conceito de pesquisa comporta a mesma subdivisão que a ciência: pesquisa pura e pesquisa aplicada"* (Simon, 2010, p. 18).

A concepção de Ryad Simon é interessante uma vez que pressupõe uma sobreposição

entre a psicoterapia e a pesquisa, pois no momento em que o terapeuta busca compreender e contextualizar a queixa, ele formula hipóteses sobre os prováveis fatores que interferem na patologia. Por outro lado, quando utiliza testes psicológicos, o psicólogo está verificando as suas hipóteses, que irão possibilitar um planejamento da psicoterapia e que irá atender às necessidades do paciente.

A psicoterapia é uma prática historicamente relevante em nosso país, razão pela qual o Conselho Federal de Psicologia instituiu 2009 como o Ano da Psicoterapia. As discussões sobre essa temática tiveram como finalidade empreender esforços na qualificação do campo, uma vez em 2004, 42% dos psicólogos dedicavam-se à psicoterapia (CFP, 2009).

Por meio de consulta à Biblioteca Virtual de Saúde (BVS) das teses indexadas utilizando na busca as palavras-chave "pesquisa e psicoterapia", no período de 1989 a 2012, foram encontradas 14 Dissertações e Teses defendidas na Universidade de São Paulo abordando o tema "*pesquisa em psicoterapia"*, entre estas, cinco utilizaram como instrumento de avaliação da psicoterapia a Escala Diagnóstica Adaptativa Operacionalizada (EDAO), desenvolvida por Simon em 1983.

Ao se considerar o modelo de psicoterapia, três estudos abordaram o modelo da Psicoterapia Breve (Cerezetti, 2002; Dantas, 2007; Yoshida, 1989), três a Psicoterapia Breve Operacionalizada (Alves, 2001; Gebara, 2009; Younes, 2011), dois a Psicoterapia Analítica Funcional (Oshiro, 2011; Taccola, 2007), um a Psicoterapia Analítico-Comportamental (Zamignani, 2008), dois a Psicoterapia Analítica Comportamental Infantil (Del Prette, 2006; 2011), um o Tratamento Padronizado e Individualizado no Transtorno do Pânico (Yano, 2003), um realizou uma Revisão Sistemática e Metanálise sobre a efetividade da psicoterapia de grupo nos casos de disfunção erétil (Melnik, 2006), e um a psicoterapia Junguiana de grupo (Duran, 1997).

O número de trabalhos relacionados a esta temática mostra o aumento de interesse pela pesquisa dos processos e resultados da psicoterapia. Segundo Yoshida, Santeiro, Santeiro e Rocha (2005), apesar de quantitativamente inferior à produção internacional, a produção nacional possui qualidade científica compatível com a de outros países.

O que se constata na literatura é o reduzido número de trabalhos envolvendo grupos de pesquisadores, de diferentes instituições e centros de pesquisas, assim como instrumentos padronizados para a realidade brasileira, que possam ser utilizados com esta finalidade. Entretanto esta realidade pode ser modificada uma vez que "*a coautoria tem sido considerada um dos itens relevantes no processo de análise da produção cientifica de uma área"* (Yoshida, et al., 2005, p. 55). De acordo com Yoshida e Rocha (2007), a avaliação da

psicoterapia tem sido valorizada pelos psicólogos brasileiros, uma vez uma vez que esta prática identifica o psicólogo como profissional da saúde.

Na literatura brasileira são mais frequentes as pesquisas de psicoterapia nas modalidades psicodinâmica breve e cognitivo-comportamental breve. Um estudo importante foi realizado por Yoshida (1989), em que a autora sintetizou as principais técnicas de psicoterapia breve psicodinâmica (PBP), assim como os respectivos critérios de indicação de cada uma delas. A autora apresentou ainda dois estudos retrospectivos de validade preditiva da Escala Diagnóstica Adaptativa Operacionalizada (EDAO).

A EDAO é um instrumento muito utilizado nas pesquisas brasileiras, a escala tem como ponto principal o critério adaptativo, que norteia a avaliação diagnóstica. Parte do pressuposto de que todo comportamento tem como objetivo a busca de adaptação, quer seja consciente ou inconsciente e necessita integrar seus diversos sistemas para ser atingida, de acordo com determinado padrão de resposta ou de comportamento. A escala avalia os seguintes setores: Afetivo Relacional (A-R), Produtividade (Pr), Sociocultural (S-C) e Orgânico (Or) (Simon, 1989).

Em seu estudo, Yoshida (1989) observou a importância dos resultados da psicoterapia. Ela procurou identificar que variáveis favorecem a obtenção de ganhos pelo paciente, partindo do seguinte questionamento: as mudanças dependem mais dos recursos adaptativos do paciente no início da terapia ou da experiência e manejo do terapeuta? Os resultados não indicaram validade preditiva da EDAO em nenhum dos dois estudos e a autora apontou como causa, possíveis falhas metodológicas, que podem ter ocasionado um viés na pesquisa.

Rocha (2002) estudou a precisão e a validade da EDAO-Revisada (EDAO-R) com pessoas idosas, que participavam de grupos da terceira idade. Seus resultados indicaram que a escala é uma medida precisa e válida da eficácia adaptativa aplicada a pessoas idosas. Outro ponto destacado foi a boa operacionalização das categorias de avaliação, possibilitando um grau satisfatório de acordo entre avaliadores independentes.

Laham, Satake, Chiba, Benute e Lucia (2009), em um estudo sobre pacientes portadores de doenças crônicas, atendidos no domicílio, descreveram a utilidade da EDAO neste contexto. Os autores concluíram que se trata de uma escala que permite a utilização de critérios mais objetivos para o psicodiagnóstico no atendimento domiciliar.

A EDAO é uma escala utilizada em psicoterapias de diferentes referenciais teóricos. Giovannetti e Sant`Anna (2005) investigaram as vivências e as necessidades psicológicas de alunos de Psicologia em relação à crise adaptativa. A escala possibilitou a avaliação dos

níveis de adaptação, colaborando de modo positivo na realização de um psicodiagnóstico preventivo e interventivo.

Os estudos mais recentes no Brasil estão pautados na busca de evidências empíricas. Yoshida (1999) estimou o grau de precisão entre avaliadores da versão revisada da Escala Diagnóstica Adaptativa Operacionalizada (EDAO-R), que pode ser utilizada como uma medida auxiliar para discriminar os pacientes com maior dificuldade de adesão ao tratamento.

O trabalho de Laloni (2001) ampliou o espectro de instrumentos disponíveis para os profissionais brasileiros. A autora estudou a adaptação, a precisão e validade da Escala de Avaliação de Sintomas – SCL-90R em uma amostra clínica. A partir dos resultados desta pesquisa, Laloni fez uma adaptação da SCL-90R. A nova escala adaptada foi denominada Escala de Avaliação de Sintomas (EAS-40), que é uma escala do tipo *Likert* de três pontos e avalia quatro dimensões: somatização, obsessividade-compulsividade, ansiedade e psicotismo. Um estudo de validade e precisão da EAS-40 em amostras não clínicas foi realizado por Yoshida e Silva (2007).

Santeiro (2005) fez um estudo sobre a produção científica sobre as psicoterapias breves com a finalidade de traçar o perfil das tendências e perspectivas na área, em dois periódicos nacionais (Estudos de Psicologia de Campinas e Jornal Brasileiro de Psiquiatria) e dois estrangeiros (Journal of Consulting and Clinical Psychology e Psychotherapy and Psychosomatics) relativos ao período de 1980 a 2002. Constatou que predominaram estudos empíricos, com adultos em atendimentos individuais, cuja principal patologia ou queixa investigada foi a depressão. Os instrumentos de avaliação psicológicos mais utilizados foram as escalas clínicas.

Muitos estudos sobre psicoterapia procuram responder o que provoca mudanças na psicoterapia? Yoshida et al. (2009) examinaram as possíveis relações entre as mudanças no padrão de relacionamento conflituoso do paciente e a estratégia terapêutica adotada pela terapeuta, assim como a “magnitude” da mudança em sintomas psicopatológicos no final do processo e em entrevistas de *follow-up* realizadas em dois momentos, 3 e 6 meses após o término do tratamento. Os resultados mostraram melhoras clinicamente significantes nos sintomas e mudança parcial do padrão central de relacionamento. Foram utilizadas intervenções expressivas (focalizadas na transferência) mais no início do processo e, intervenções mais suportivas (não focalizadas na transferência), na medida em que mudanças positivas foram observadas. Entretanto, segundo os autores, como se trata de estudo de um único caso, é necessária certa cautela quanto à generalização dos resultados.

Yoshida (2008) realizou um estudo cujo objetivo foi aprofundar o conhecimento sobre

o método de Jacobson e Truax para determinar mudanças clinicamente significativas de sintomas avaliados por escalas de autorrelato, especialmente no contexto onde não há dados normativos a serem utilizados como referência. Neste trabalho foi utilizada uma amostra de estudantes universitários, como referência da população funcional estabelecendo assim os escores de mudança e notas de corte de medidas de sintomas. No texto foi discutida a escolha do critério de mudança, assim como formas de avaliar a magnitude da mudança em um processo de psicoterapia breve. A questão que desafia o pesquisador nesta área é a de escolher procedimentos que permitam a avaliação de mudança que tenha significância clínica e não apenas significância estatística. Segundo Jacobson, et al. (1991), dois critérios devem ser considerados para que uma mudança possa ser considerada clinicamente significativa: primeiro o paciente deve evoluir de um funcionamento disfuncional para um funcional; segundo apresentar uma mudança acentuada de modo a ser possível considerar como uma mudança real e não como um resultado "positivo" devido a um erro de medida. Contudo, apesar dos limites, os critérios podem ser usados em pesquisa de avaliação de mudanças em psicoterapia breve.

A proposta de Jacobson, Follette e Revenstorf (1984) de operacionalização dos critérios para avaliação de mudança clinicamente significativa é muito importante por isto merece ser destacado. Os autores ressaltam que a nota de corte para considerar uma mudança pode variar em função da amostra utilizada como referência. Após o término do tratamento, os escores obtidos pelo paciente, devem classificá-lo dentro da faixa de dispersão de funcionamento da população não clínica, que é definida por dois desvios padrão em relação à média da amostra considerada como referencial. A aplicação desses critérios exige do psicólogo conhecimento das normas do teste utilizado na avaliação da mudança.

Os estudos sobre psicoterapia são realizados por pesquisadores de diferentes referenciais teóricos e com uma variedade de instrumentos. Prado e Meyer (2006) analisaram a relação terapêutica por meio do *Working Alliance Inventory* (WAI) em atendimentos realizados pela internet. Concluíram que a relação desenvolvida apresentou características semelhantes às descritas na literatura, quanto ao atendimento presencial empregando o mesmo instrumento. A aliança de trabalho é uma das variáveis importantes na interação da díade terapeuta-paciente, pois influencia no processo e nos resultados da psicoterapia.

As revisões literatura com a finalidade de analisar a produção científica são importantes para os pesquisadores, uma vez que compilam informações de diferentes pesquisas sobre uma determinada temática. Jung, Fillipon, Nunes e Eizirik (2006) publicaram um histórico sobre a pesquisa de resultados especificamente em psicoterapia psicanalítica de

longa duração, no período de 1990 a 2006. Segundo Wallerstein (apud Jung, et al., 2006), as pesquisas da primeira geração (1917 a 1960) foram caracterizadas pelas contagens estatísticas de resultados em diferentes categorias de pacientes. Na segunda geração, as pesquisas utilizavam medidas desenvolvidas para avaliar o pré/pós-tratamento e preditores para os resultados. Na terceira geração, foram incluídas as medidas de resultados e estudos de *follow up.* A quarta geração, iniciada em 1980, e ainda em desenvolvimento, envolve novas medidas de resultados para conhecer de modo mais preciso os mecanismos de mudança.

Historicamente as pesquisas em outros países datam do início do século passado, para identificar que variáveis contribuem para o bom andamento da psicoterapia e permitem aos pacientes obterem ganhos com o processo terapêutico. Segundo Strupp e Howard (apud Neno, 2005), a partir de 1920 foram produzidos os primeiros estudos que reuniram dados mais sistematizados dos resultados de psicoterapias, mas o número de pacientes era reduzido e o processo de psicoterapia não era descrito de modo mais minucioso.

Em 1952 Eysenck publicou um artigo de revisão da literatura sobre resultados de psicoterapia, concluindo que os dados obtidos não foram suficientes para sustentar a hipótese de que a psicoterapia facilita a recuperação do paciente. Este artigo gerou um movimento que levou muitos profissionais a se comprometeram a avaliar processo e resultados dos tratamentos. Strupp, em 1963, em uma réplica às críticas de Eysenck, escreveu sobre os objetivos do tratamento psicanalítico, destinado a tornar os pacientes indivíduos livre das inibições incapacitantes, dos sintomas, com capacidade de autocompreensão e aceitação tanto de seus pontos fortes como fracos e estabelecer e manter relacionamentos interpessoais duradouros (Lambert, 2013).

De acordo com Hilsenroth (2013), foi Strupp quem propôs uma questão mais ampla sobre como o campo da psicoterapia poderia avançar para avaliar o processo de psicoterapia e os resultados. Grande parte das pesquisas que surgiram a partir daí se originaram das ideias de Strupp, influenciando o foco das pesquisas até os dias de hoje. Neste primeiro artigo *The Outcome problem in psychotherapy revisited publicado em 1963,* ele defendia a importância da avaliar os resultados de modo global, com especial destaque para os aspectos do funcionamento interpessoal e intrapessoal até as pesquisas de casos individuais, utilização de diferentes escalas de avaliação, bem como de relatórios de atendimento se referiam ao diálogo terapêutico. Também foram as ideias de Strupp que deram base para o desenvolvimento do método de significância clínica dos tratamentos, e antecipou a necessidade de examinar as características do paciente e terapeuta, assim como os efeitos mediadores das variáveis de processo nos resultados de psicoterapia.

Os efeitos da psicoterapia foram investigados por meio de uma variedade de estudos e ao longo de um período que remonta aos anos 30. Os pesquisadores têm realizados esforços redobrados para melhorar consideravelmente os resultados. Segundo Lambert (2007) uma forma de melhorar os resultados da psicoterapia é acompanhar o progresso do paciente, uma prática que precisa ser assumida institucionalmente e fornecer estas informações ao terapeuta, a fim de orientar o tratamento em curso, e especial atenção para o paciente, que não está tendo uma resposta favorável à psicoterapia.

A avaliação dos resultados da psicoterapia durante muito tempo ficou a cargo da experiência pessoal pessoais do terapeuta, o que não garantia uma avaliação isenta da subjetividade do profissional. Segundo Luborsky (1994) o terapeuta tem uma relação muito próxima pela facilitação do vínculo formado na relação terapêutica, assim a visão do paciente prediz melhor os resultados.

Barlow, Boswell e Thompson-Hollands (2013) fizeram uma revisão sobre pesquisa em psicoterapia ao longo dos últimos 50 anos, desde as primeiras concepções de Strupp e Eysenck assim como a realização de muitas de suas próprias metas para a pesquisa. Concluíram com uma previsão para o futuro com base no conhecimento acumulado de psicoterapia e psicopatologia e, no contexto de que o conhecimento, o tipo de programas de pesquisa ideais que serão necessários.

No número comemorativo foi republicado o artigo original de 1963, Strupp (2013) escreveu que observações clínicas amplamente documentadas em vários artigos mostram que pacientes comenta que observações clínicas amplamente documentadas em vários artigos mostram que pacientes, ao procurar ajuda em função de dificuldades pessoais, se beneficiam da relação interpessoal com um profissional. Hoje não é possível mais argumentar contra esta ideia, pois a psicoterapia traz benefícios para os pacientes. Atualmente os objetivos das pesquisas têm como finalidade compreender mais profundamente a interação terapeuta-paciente Cada vez mais os pesquisadores estão interessados em questões importantes, tais como: Que tipos de mudanças de personalidade ocorrem como resultados de determinados tipos de intervenções? Como variáveis do terapeuta e paciente influenciam a interação terapêutica? Muitas perguntas ainda não foram respondidas por causa da dificuldade em medir com maior precisão os resultados. Avaliar a natureza das perturbações do paciente antes de iniciar a terapia, é um dos aspectos mais importantes nos estudos de avaliação de resultados, que considera a melhora do paciente. Sob que condições específicas a psicoterapia é mais eficaz? Segundo Strupp esta é uma questão legítima para futuras pesquisas.

No âmbito internacional observa-se que os estudos são realizados por grupos de

pesquisadores de diferentes centros em todo o mundo. A *Society of Psychotherapy Research* (SPR) agrega participantes de diferentes países, para a discussão das especificidades do tratamento, dos delineamentos e resultados dos trabalhos. No *43rd Annual International Meeting* da SPR (2012) foram apresentados 15 estudos que utilizaram a escala OQ-45.2 para avaliar processo e/ou resultados de psicoterapias. O avanço das pesquisas exige o uso medidas válidas e precisas, para avaliação de processos e resultados, que possam ser aplicadas em diferentes modelos terapêuticos.

Instrumentos para avaliação da psicoterapia

Segundo Urbina (2007), os testes podem ser classificados em três categorias: a primeira se refere ao processo de tomada de decisões a respeito do(s) indivíduo(s). A segunda categoria é voltada para a pesquisa científica de fenômenos psicológicos e diferenças individuais. E o uso mais recente se refere ao contexto da avaliação de processos e resultados terapêuticos.

De acordo com Anastasi (1977, p. 521), "*os testes de personalidade são instrumentos para a mensuração de características emocionais, de motivação, interpessoais e de comportamento"*. A autora classifica os principais tipos de testes psicológicos para avaliar a personalidade como: inventários de personalidade, medidas de interesse e de atitudes e técnicas projetivas.

O teste psicológico é um procedimento sistematizado para a obtenção de amostras de comportamentos relevantes do funcionamento cognitivo ou afetivo, com a finalidade de avaliar amostras destes comportamentos de acordo com determinados padrões. Os testes são descritos como padronizados, devido à necessidade de objetividade no processo de avaliação psicológica. A uniformidade de procedimentos engloba desde a aplicação, as circunstâncias de sua administração, a avaliação e a interpretação dos resultados (Urbina, 2007).

Segundo Cronbach (1996) as escalas de autorrelato têm sido utilizadas na avaliação clínica de psicopatologias. Uma das primeiras escalas publicadas foi a *Minnesota Multiphasic Personality Inventory* (MMPI) de Hathaway e McKinley, cujo objetivo é avaliar diferentes aspectos da personalidade. Na década de 60 surgiram os inventários para sintomas unidimensionais, como os Inventários de Depressão, Ansiedade, Desesperança e Ideação Suicida de Beck, desenvolvidos por Beck, Ward, Mendelson, Mock e Erbaugh (Gorenstein, Pang, Argimon & Werlang, 2011), o Inventário de Ansiedade-Traço, criado por Spielberger, Gorsuch e Lushene (Biaggio & Natalício, 1979) e a Escala de Personalidade de Comrey, elaborada por Comrey (Costa, 1997).

Entre os inventários multidimensionais destacam-se o questionário de 16 Fatores de Personalidade – 16PF, muito utilizado no Brasil e elaborado por Raymond Cattell e o Questionário de Saúde Geral de Goldberg (QSG) criado por Goldberg e concebido para avaliar a saúde mental. Segundo os autores da adaptação brasileira do QSG (Pasquali, Gouveia, Andriola, Miranda & Ramos, 2011) o conceito de saúde mental é vago e pode ser mais bem definido em termos dos desvios comportamentais, tendo como referência a população geral. O instrumento é composto por 60 itens de sintomas psiquiátricos não psicóticos. É uma escala *Likert* de quatro pontos, organizada para avaliar cinco fatores: Stress psíquico, Desejo de morte, Desconfiança no próprio desempenho, Distúrbio do sono, Distúrbios psicossomáticos e o Fator Geral denominado Saúde Geral. Quanto maior a pontuação no escore do Fator Geral, maior será o nível de distúrbios psiquiátricos. Para a adaptação brasileira do QSG, a amostra foi composta de 902 participantes, de São Paulo e Distrito Federal, destes 314 do sexo masculino e 588 do sexo feminino, com idade entre 19 e 70 anos, e com predomínio de estudantes universitário.

As escalas de autorrelato como medida de psicopatologia apresentam características que as recomendam nesta área e, nos últimos anos, foram desenvolvidas as escalas para avaliação de resultados de processos psicoterápicos. O Inventário de Relacionamento, desenvolvido por Barret-Lennard para avaliar as condições oferecidas pelo terapeuta, tem como referência teórica a terapia centrada no cliente de Rogers (Ramos, 2008). Este foi um dos primeiros instrumentos de avaliação da aliança terapêutica usado em estudos de avaliação de resultados de psicoterapia, o qual mostrou que a percepção dos pacientes sobre o relacionamento se modifica durante o curso do tratamento. O Relato da Sessão Terapêutica de Orlinsky e Howard foi desenvolvido para avaliar a experiência do paciente na psicoterapia, abordando quatro aspectos da experiência terapêutica: sentimentos, troca, diálogo e relacionamento entre pacientes e terapeutas (Orlinsky & Horward, 1986).

Uma das primeiras pesquisas de tradução de instrumento para avaliação de aliança terapêutica no Brasil foi realizado por Marcolino e Iacoponi (2001), a Escala de Aliança Psicoterápica da Califórnia na versão paciente; os autores relatam associação positiva entre o grau de aliança terapêutica e os resultados da psicoterapia.

Morgan, Luborsky, Crits-Christoph e Solomon (1982), no Projeto de Psicoterapia Penn, relataram progressos na identificação dos fatores preditivos de resultados de psicoterapia, utilizando o *Penn Helping Alliance Counting Signs Method.* Identificaram dois tipos de aliança, no primeiro tipo, o paciente percebe o terapeuta como o maior responsável pelo avanço da terapia e, no segundo, percebe o tratamento como um processo colaborativo

em que ambos, terapeuta e paciente, trabalham em conjunto. Os pacientes que percebem a psicoterapia como um processo em que a responsabilidade é compartilhada entre ele e o terapeuta, obtiveram os melhores resultados (Luborsky, Mclelan, Woody, O' Brien & Auerback, 1985).

O Inventário de Aliança de Trabalho (*Working Alliance Inventory - WAI*) foi desenvolvido por Horvath e Greenberg como medida de aliança terapêutica em três versões: terapeuta, paciente e observador (Horvath & Greenberg, 1994). Segundo Peres (2009), a aliança terapêutica é uma variável central do processo psicoterapêutico, que pode ser entendida como um critério de indicação para psicoterapia e também como um fator preditivo de sua evolução.

Características psicométricas dos testes psicológicos

Para Van Kolck (1974), as qualidades consideradas essenciais de um instrumento de avaliação são: validade, precisão ou fidedignidade, padronização e aferição. Assim como a homogeneidade e a sensibilidade discriminativa.

A precisão do teste é "*a consistência dos resultados obtidos pelos mesmos indivíduos em diferentes oportunidades ou com diferentes conjuntos de itens equivalentes"* (Anastasi, 1977, p.84). O conceito de precisão está relacionado ao erro de mensuração de um único resultado e indica quanto as diferenças nos resultados do teste se devem a diferenças reais e quanto podem ser devidas a erros casuais. Para Cronbach (1996), a precisão se refere à proporção da variância atribuída ao escore verdadeiro da variável latente. Os instrumentos de avaliação psicológica devem fornecer medidas confiáveis, de modo que se obtenham os mesmos resultados ao repetir a medida em condições semelhantes (Bunchaft & Cavas, 2002).

Segundo Pasquali (2003), são utilizados vários termos para definir essa característica essencial que um teste deve possuir. Os principais são precisão, fidedignidade e confiabilidade. Existem diferentes métodos para obter a precisão de um teste: teste-reteste, formas alternativas e consistência interna, além da precisão do avaliador.

A Precisão Teste-Reteste consiste em fazer uma nova aplicação ou repetição de um teste e correlacionar os resultados obtidos pelos mesmos sujeitos nas duas ocasiões. É importante considerar o intervalo de tempo entre as duas aplicações, pois o coeficiente de correlação diminui progressivamente, quando o intervalo de tempo aumenta. Se o intervalo for muito curto pode haver a influência da memória ou da aprendizagem, dependendo da variável avaliada (Anastasi, 1977; Pasquali, 2003). Este método avalia principalmente a

estabilidade temporal.

A Precisão de Formas Alternativas consiste na aplicação de duas formas equivalentes ou paralelas do mesmo teste, cujos resultados obtidos nas duas formas são correlacionados. A equivalência entre as duas formas do teste é obtida pela análise do conteúdo dos itens, uma vez que devem ter o mesmo número de itens, ser apresentados da mesma forma, abranger o mesmo conteúdo e ter o mesmo nível de dificuldade (Anastasi, 1977; Pasquali, 2003).

A Precisão da Consistência ou Coerência Interna pode ser obtida por várias técnicas, sendo que as principais são: a precisão das metades, Kuder-Richardson e Alfa de Cronbach (Pasquali, 2001). Nesse caso é necessária apenas uma única aplicação do teste.

A Precisão das Metades é obtida pela divisão do teste em duas metades comparáveis ou equivalentes, sendo a mais comum a divisão em itens pares e impares. Correlacionam-se os totais das duas metades e depois é aplicada a correção pela fórmula de Spearman-Brown.

A Precisão de Kuder-Richardson, que verifica a coerência entre os itens, é obtida por uma fórmula e seu resultado corresponde à média de todos os coeficientes de metades de todas as divisões possíveis do teste.

O Alfa de Cronbach é uma extensão da fórmula de Kuder-Richardson, que só pode ser aplicada, quando a variável é dicotômica, com duas possibilidades de resposta, isto é, certo ou errado (zero ou um). Quando a resposta ao item tiver mais de duas alternativas, usa-se o coeficiente Alfa.

Validade

A validade de um teste está relacionada "*àquilo que o teste mede e quão bem ele o faz"* (Anastasi e Urbina, 2000, p.105). A validação é um processo por meio do qual as evidências de validade são coletadas. A validade de um teste pode aumentar à medida que as evidências empíricas para interpretação dos resultados do teste se acumulam e demonstram que os fundamentos teóricos do teste são confirmados ao avaliar as evidências disponíveis no contexto em que este é usado (Urbina, 2007).

Segundo Anastasi (1977) a validade pode ser classificada em três tipos: de conteúdo, de critério e de constructo. A Validade de Conteúdo pressupõe um exame sistemático do teste de modo a assegurar que os itens representem aspectos fundamentais do comportamento a ser medido. A Validade de Critério refere-se à validação de um instrumento em relação a um critério, por exemplo, nos casos em que se quer predizer o desempenho futuro, é feita a correlação entre os resultados obtidos no teste e os resultados dos mesmos indivíduos em um critério. A Validade de Constructo de um teste é derivada da sua fundamentação teórica,

busca a demonstração de que o instrumento de medida avalia determinado constructo.

> A validade, assim como a fidedignidade, não é uma qualidade que caracteriza abstratamente os testes ou qualquer teste específico ou seus dados. Mais do que isso, a validade é uma questão de julgamento que diz respeito aos escores do teste, como são empregados para um determinado objetivo em um dado contexto. Por isso, o processo de validação é semelhante à testagem de hipóteses (Urbina, 2007, p. 156).

Na atualidade, "*a validade de constructo passa a ser reconhecida como o conceito de validade fundamental e totalmente inclusivo"* (Anastasi & Urbina, 2000, p.108), pois inclui os outros tipos de validade que eram considerados anteriormente.

OUTCOME QUESTIONNAIRE – OQ-45.2

O OQ-45.2 é um dos instrumentos mais utilizados para avaliação de psicoterapia, segundo a literatura internacional, o qual se mostra bastante útil para atingir múltiplos objetivos na área de avaliação de processos e resultados de psicoterapias. Segundo Lambert et al. (2004), o instrumento foi traduzido em diversas línguas: espanhol, alemão, francês, holandês, sueco, norueguês, hebraico, italiano e russo. As versões em alemão e em espanhol foram validadas.

Diversas pesquisas realizadas em outros países verificaram suas propriedades psicométricas (Abanishe, 2008; Lambert et al., 1996; Lo Coco, et al., 2008; Kim et al. 2007; Mueller, Lambert & Burlingame, 1998). Outros estudos utilizaram OQ-45.2 no campo de avaliação de processos psicoterápicos (Beckstead, et al., 2003; Cisneros, 2010; Machado, Klein e Farate, 2005; Martínez et al., 2004; Okiishi, Lambert, Eggett, Nielsen & Dayton, 2006).

A versão utilizada para esta pesquisa é resultado de uma tradução e adaptação cultural do OQ-45.2 realizada por Carvalho e Rocha (2009), após a autorização dos autores para tradução, adaptação cultural e equivalência semântica para o uso no Brasil.

Carvalho e Rocha (2009) traduziram e adaptaram o OQ-45.2 para o contexto brasileiro. O processo ocorreu em cinco etapas: tradução (teve como base duas versões, a original em inglês e a portuguesa), versão síntese, *backtranslation*, revisão por especialistas e pré-teste. Utilizaram o guia de Beaton e colaboradores, considerado uma referência para os estudos de adaptação transcultural. Após a tradução da primeira versão brasileira, na etapa II as três versões (original, portuguesa e a versão brasileira) foram comparadas, discutidas e como resultado obteve-se uma versão síntese. Na etapa III (*backtranslantion*), a versão síntese foi encaminhada a um tradutor profissional para realizar a tradução para a língua original do instrumento, isto é, uma versão em inglês a partir da versão traduzida para o português. Na etapa IV as versões produzidas nas etapas anteriores foram revisadas e comparadas pelos dois pesquisadores, com o objetivo de verificar a equivalência entre elas, do ponto de vista semântico, idiomático, experiencial e conceitual. Na etapa final foi realizado um pré-teste, a versão da escala foi aplicada em uma amostra com diferentes graus de escolaridade e posteriormente os participantes foram questionados quanto à compreensão dos itens. Pela clareza e objetividade como os itens foram construídos na escala original, não foram encontradas dificuldades para a equivalência semântica. Os autores consideraram os procedimentos adotados para elaboração da versão brasileira adequados e suficientes para

futuras pesquisas das propriedades psicométricas da versão brasileira da escala.

O OQ-45 é um instrumento construído para avaliar mudanças no paciente durante o tratamento psicológico, que pode ser utilizado ao longo ou ao final do processo para medir os resultados da psicoterapia. Segundo a concepção de Lambert (1983), o progresso do paciente é avaliado por meio de três aspectos da sua vida: o desconforto subjetivo (funcionamento intrapsíquico), as relações interpessoais e o desempenho do papel social. Estas áreas de funcionamento sugerem um continuum dos sentimentos e sensações percebidas do seu estado interior, o que experencia no seu íntimo, como se relaciona com as pessoas significativas da sua vida e como lida com tarefas relacionadas à produtividade, por exemplo, o trabalho, a escola ou qualquer outra atividade, inclusive o lazer. Além disso, o instrumento foi concebido para estabelecer uma linha de base, que pode ser utilizada como referência na tomada de decisões em momentos importantes do tratamento psicoterápico (Lambert et al., 2004).

O OQ-45.2 foi concebido por Lambert, Lunnen, Umphress, Hansen e Burlingame em 1994 (Lambert & Finch, 1999). Não foi idealizado para o diagnóstico de pacientes, pois esta tarefa pode ser realizada com instrumentos desenvolvidos para esta finalidade e para isso citam o MMPI-2. O OQ-45.2 foi concebido para o uso em três níveis distintos de aplicação: 1) mensurar a dimensão do sofrimento psíquico atual, 2) como uma escala de avaliação de resultados deve ser aplicada previamente e após as intervenções no tratamento ou para monitorar a resposta do paciente ao tratamento e, 3) acompanhar por meio de um instrumento informatizado a tomada de decisões para melhorar a qualidade do tratamento.

Nos E.U.A. a *Human Affairs International* (HAI), empresa estatal de assistência à saúde dispõe de um serviço informatizado denominado *The Adult Clinical Support Tool, OQ-ASC (Assessment For Signal Clients),* que permite aos profissionais tomarem decisões quanto ao tratamento com base no progresso do paciente em vez de se basearem em um protocolo de atendimento fixo ou políticas obrigatórias em relação ao tempo de tratamento. A partir da pontuação obtida no OQ-45.2 os pacientes são classificados em categorias, por exemplo, alta (escore total de 85 ou acima), médio (entre 64-84) ou baixo (63 ou menor). Além dos escores do OQ-45.2, são combinadas informações sobre o histórico de tratamento psicológico, cuidados hospitalares recentes, a motivação para a psicoterapia e diagnóstico, gerando relatórios feitos por computadores que fornecem sugestões aos profissionais e gestores para o planejamento dos tratamentos ou encaminhamento.

Com base na avaliação inicial do paciente, alguns pacientes podem receber encaminhamento para psicoterapias breves (1 a 8 sessões), outros são encaminhados para tratamento de longo prazo, uso de medicamentos, intervenção nos casos de abuso de

substâncias ou psicoterapia de grupo. Além disso, quando ocorre uma variação nos escores do OQ-45.2, indicando que o paciente apresenta uma piora em termos de sofrimento subjetivo ao invés de uma melhora, a comparação dos resultados entre o pré e o pós teste permite ao profissional modificar as estratégias de tratamento para prevenir ou reduzir abandono do tratamento psicológico.

A escala foi desenvolvida para superar as limitações de outras escalas de medida de resultados. Especificamente, o baixo custo, além da sensibilidade às alterações sintomatológicas em curtos intervalos de tempo, sem que com isso comprometa a sua confiabilidade e validade. Também abrange uma ampla variedade de síndromes e transtornos mentais prevalentes em adultos, entre os quais, doenças relacionadas ao estresse.

Os estudos de Lambert, Hansen e Finch (2001) apontam incremento nos resultados terapêuticos, ou seja, o maior montante de mudanças obtidas em menor período de tempo. Além destas funções, avaliar progresso, mudanças e incrementar os resultados, o OQ-45 também é considerado pelos próprios autores como apropriado para medida de funcionamento global, o que permite aferir o estado de funcionamento inicial do indivíduo, servindo como parâmetro para eventuais mudanças posteriores. Este instrumento pode fornecer uma avaliação fidedigna de diversos aspectos do nível de ajustamento e perturbações psicossocial dos indivíduos. Por suas características, ele poderia substituir adequadamente vários outros instrumentos e otimizar o tempo despendido na avaliação do atendimento psicológico.

A versão portuguesa do OQ-45.2 foi elaborada por Machado e Klein em 2002, entretanto não disponibilizaram dados relativos às características psicométricas do instrumento na população portuguesa. Machado, Klein e Farate (2005) elaboraram um projeto com o objetivo de averiguar e determinar variáveis que contribuem para o sucesso e/ou insucesso terapêutico, determinar a eficácia e eficiência dos serviços prestados no Centro de Atendimento à Toxicodependência Oriental do Porto. Desenvolveram um protocolo padrão de avaliação dos pacientes, buscando aumentar a validade externa relativa aos conhecimentos teóricos e práticos no tratamento do consumo e dependência de substâncias psicoativas.

De acordo com Lambert et al. (2004), a seleção dos itens para o *OQ-45.2* foi determinada por uma série de considerações. Inicialmente foram selecionadas questões que abordam disfunções ou problemas mais comuns em uma ampla variedade de distúrbios. A seguir, itens que abordam os sintomas, com maior probabilidade de ocorrência, independente da particularidade dos problemas. E, na terceira etapa, os itens para avaliar características

pessoais e sociais relevantes sobre a qualidade de vida. No final, o número de itens foi reduzido com a finalidade de facilitar o uso do teste no dia a dia da prática clínica e para permitir que ele fosse mais facilmente aceito pelos pacientes em repetidas reaplicações, bem como para fornecer dados clínicos para situações de tomada de decisão ao longo do processo terapêutico.

O teste apresenta três subescalas, que foram obtidas como resultado de uma ampla análise fatorial. O desenvolvimento dos itens foi baseado na concepção de que três grandes áreas de conteúdo são de importância crucial para medir o estado do paciente e os resultados da psicoterapia. Essas três áreas são resumidas por Lambert e Hill (1994), refletindo "a necessidade" de avaliar mudanças no cliente, tais como a diminuição ou desaparecimento do desconforto subjetivo, nas suas relações mais íntimas (relacionamentos conjugais, familiares e de amizades) e na sua participação na comunidade e desempenho dos papéis sociais, como trabalho e atividades de lazer. Os itens do teste foram agrupados em três fatores denominados: Desconforto Subjetivo, Relações Interpessoais e Papel Social.

Desconforto Subjetivo (SD)

A escolha dos itens relacionados ao domínio Desconforto Subjetivo (SD) se originou de um estudo do *National Institute of Mental Health* (NIMH), no qual foram identificados os tipos de doenças mentais de maior prevalência em cinco áreas dos Estados Unidos (EUA). Foi realizada uma pesquisa documental pelo NIMH, nos arquivos de uma empresa de seguros, sobre os transtornos mentais, com a finalidade de investigar a frequência dos códigos de diagnóstico, segundo o DSM III-R.

De acordo com Lambert e Finch (1999) o estudo epidemiológico do NIMH de 1988, com 18.571 sujeitos, revelou que 15,4% da população americana com mais de 18 anos preenchiam pelo menos um dos critérios de transtorno mental. Entre os quais, 12% preencheram os critérios de diagnóstico de ansiedade ou de transtorno afetivo.

A partir da consulta ao banco de dados da *Human Affairs International (HAI),* com os códigos dos diagnósticos de 2.145 pacientes, os dados indicaram que quase um terço dos diagnósticos referia-se a algum tipo de transtorno afetivo. Os outros diagnósticos envolviam algum tipo de transtorno de ansiedade, incluindo transtorno de estresse pós-traumático. Estes dados sugeriram que os sintomas intrapsíquicos mais comuns nos EUA seriam depressão e ansiedade (Lambert et al., 2004).

Segundo Feldman (apud Lambert et al; 2004) os sintomas de ansiedade e os de depressão não podem ser separados facilmente, uma vez que tendem a ocorrer em uma ampla

variedade de diagnósticos de diversos transtornos. Os autores do OQ-45.2 não fizeram nenhuma tentativa de delinear escalas separadas de ansiedade e depressão, desenvolveram o domínio Desconforto Subjetivo com base em itens relativos à ansiedade e depressão. Acrescentaram ainda itens referentes ao uso de substâncias psicoativas, por ter sido o problema diagnosticado com maior prevalência. A escala de desconforto subjetivo busca identificar aspectos psicopatológicos referentes a transtornos de humor, transtornos de ansiedade e transtorno de abuso de substâncias.

Relações Interpessoais (IR)

O OQ-45.2 foi elaborado com itens que avaliam a satisfação geral e os problemas nas relações interpessoais. Segundo Lambert et al. (2004), o desenvolvimento dos itens do domínio Relações Interpessoais (RI) foi baseado nas pesquisas de Andrews e Whitney (1974); Beiser (1983); Blau (1977); Deiner (1984) e Veit e Ware (1983) sobre o grau de satisfação individual e qualidade de vida. Os resultados destas pesquisas sugerem que as pessoas consideram essenciais as relacões interpessoais positivas para a sensação de felicidade. As pesquisas de Horowitz et al., mostraram que os conflitos de natureza interpessoal são os problemas mais frequentemente relatados pelas pessoas que buscam psicoterapia. Também foram relevantes as pesquisas de Deiner (1984) e de Zautra (1983), que enfatizam que as relações interpessoais mais íntimas têm uma contribuição fundamental para a sensação de bem-estar. Lambert et al. (2004) fazem referência ainda ao estudo de Horowitz, et al. (1981), no qual os autores afirmam que os problemas nas relações interpessoais são claramente associados à angústia intrapsíquica, como causa direta ou como resultante de transtornos psicopatológicos.

Consequentemente em função dos dados das pesquisas citadas, itens relacionados aos afetos nas relaçoes de amizade, vida familiar e casamento foram incluídos na avaliação, assim como itens que se referem ao grau de atrito, conflito, isolamento, inadequação e rompimento nas relações interpessoais. Os itens que compõem o domínio Relações Interpessoais foram desenvolvidos com base nas pesquisas científicas sobre terapia de casal e familiar, bem como na investigação da frequência dos problemas interpessoais mais comumente relatados por pacientes em psicoterapia. Desta forma, o domínio relacionamento interpessoal visa avaliar os relacionamentos dos individuos, mais especificamente os conjugais, familiares e de amizades.

Papel Social (SR)

Os itens que compõem o domínio desempenho do Papel Social foram elaborados com base na frequência dos relatos de pacientes sobre insatisfação, conflito, angústia e sentimentos de inadequação, relacionados ao trabalho, desempenho do papel familiar e formas de lazer. Os autores do OQ-45 basearam-se nas pesquisas de Frisch, Cornell, Villanueva e Retzlaff (1992), que sugerem que os sintomas e problemas psicológicos afetam a capacidade de trabalhar, amar e divertir-se e, desse modo, influenciam o desempenho do papel social. Os dados confirmados, tanto pelas pesquisas, como pela observação de pessoas que desenvolvem sintomas mostram que comumente elas têm a sua vida pessoal e profissional afetada. Lambert et al. (2004) também relatam que Kopta, Howard, Lowrey e Beutler (1994) constataram que esses sintomas podem existir independentemente de sintomas psicológicos e angústia. Portanto foram incluídos itens para avaliar o desempenho do papel social no trabalho, atividades voluntárias e lazer, uma vez que o grau de satisfação nessas duas áreas apresenta alta correlação com o nível geral de satisfação com a vida.

De modo geral, o OQ-45.2 é um instrumento a ser usado em triagens rápidas e é uma escala de resultados, que busca avaliar as experiências subjetivas e de como a pessoa funciona no mundo.

APLICAÇÃO E AVALIAÇÃO DO OQ-45.2

A descrição relativa à aplicação, avaliação e de padronização americana foi extraída de Lambert et al. (2004). Na aplicação é importante incentivar o participante a responder de modo sincero e completar todas as questões com consciência e honestidade. Deve-se lembrar que o comportamento do examinador pode interferir nas respostas dos participantes. Assim atitudes negativas do examinador podem comprometer a validade do teste, bem como por motivos pessoais, os respondentes podem fornecer uma imagem menos sincera de si mesmos.

Em circunstâncias normais, o tempo médio para completar a escala é de cinco minutos. Pessoas especialmente cautelosas podem gastar em torno de 15 minutos, enquanto outros mais indecisos podem responder o teste em períodos mais longos.

O OQ-45.2 é um instrumento de autorrelato, autoaplicável e não requer nenhuma instrução além das impressas na folha de respostas. Possui 45 itens cujas respostas são dadas numa escala tipo *Likert* de cinco pontos que variam entre “nunca” (zero ponto), “raramente” (um ponto), “as vezes” (dois pontos), “frequentemente” (três pontos) e “quase sempre” (quatro pontos). Entretanto, é necessária atenção especial aos nove itens cuja pontuação é apresentada no sentido inverso (1, 12, 13, 20, 21, 24, 31, 37 e 43). A pontuação total é obtida pela soma dos valores dos 45 itens, podendo oscilar entre zero e 180, sendo que, quanto maior a pontuação, mais comprometido está o paciente.

No caso de omissão de respostas, o valor da resposta omitida é substituida pela média dos pontos dos outros itens respondidos na mesma subescala, arredondando para o número inteiro mais próximo.

A pontuação da subescala SD (Desconforto Subjetivo) é obtida pela soma dos valores dos 25 itens, 2, 3, 5, 6, 8, 9, 10, 11, 13, 15, 22, 23, 24, 25, 27, 29, 31, 33, 34, 35, 36, 40, 41, 42 e 45, variando de zero a 100.

Para a subescala IR (Relacionamentos Interpessoais) a pontuação é calculada pela soma dos valores dos 11 itens 1, 7, 16, 17, 18, 19, 20, 26, 30, 37 e 43, variando de zero a 44 pontos.

A pontuação da subescala SR (Papel Social) é calculada somando-se os valores dos 9 itens 4, 12, 14, 21, 28, 32, 38, 39 e 44, variando de zero a 36 pontos.

Para interpretação do Escore Total do OQ 45.2, os autores consideram que quanto maior for a pontuação, maior será a percepção do paciente sobre a existência de desconforto subjetivo. A distribuição na amostra americana das notas de corte para o escore total do OQ

45.2 e as subescalas são: Escore Total = 63, Desconforto Subjetivo = 36, Relações Interpessoais = 15 e Papel Social = 12, na amostra americana (Lambert et al , 2004). Na amostra holandesa, as distribuições das notas de corte são: Escore Total= 54; Desconforto Subjetivo = 33; Relações Interpessoais = 12; Papel Social = 11. (Jong et al, 2007).

Padronização Americana

A padronização do OQ-45.2 nos EUA abrangeu a coleta de dados em diferentes localidades geográficas e grupos diversos. Para obtenção de confiabilidade o teste foi aplicado a estudantes universitários de Utah, Idaho, e Ohio. A aplicação durou em média 20 minutos. Três semanas após a primeira aplicação, o teste foi reaplicado e os coeficientes de confiabilidade pelo reteste foram calculados pela correlação de Pearson. Também foram obtidos dados da estabilidade temporal em testes aplicados semanalmente.

Os dados normativos relativos à comunidade foram coletados em diferentes localidades. Uma amostra de 208 sujeitos foi obtida em Utah, sendo os participantes selecionados aleatoriamente na lista telefônica e contatados por telefone. Os adultos presentes na residência foram convidados a responder o questionário, a fim de colaborar para que o pesquisador pudesse entender melhor como as pessoas responderiam ao instrumento. Os testes foram enviados aos que concordaram em participar, junto com o termo de consentimento e um envelope selado para o retorno. Após uma semana, foram contatados novamente por telefone, para verificar se já haviam completado a escala. Os que não tinham respondido foram incentivados a preencher o instrumento. As respostas foram anônimas para incentivar a maior franqueza.

Dados normativos adicionais para a padronização foram coletados em uma grande companhia seguradora de porte nacional, com 800 funcionários, que autorizou a aplicação do OQ-45.2. Foi enviada uma carta convite com a assinatura do autor principal do estudo a cada um dos funcionários. O objetivo do teste era explicado e os funcionários solicitados a completar o OQ-45 e devolvê-lo em um envelope selado. A participação foi voluntária e os participantes foram instruídos a não fornecer nenhuma informação que pudesse identificá-los. Dos 800 testes expedidos, 365 (45%) foram respondidos. Esse mesmo procedimento foi adotado em Ohio, em uma grande variedade de empresas. Os dados foram analisados por meio de análise de variância.

Os dados das amostras clínicas foram coletados pelas recepcionistas da clínica, que aplicaram o OQ-45.2 antes da primeira sessão e em todas as sessões subseqüentes. Foram incluídos ainda dados do arquivo de um centro de atendimento de uma grande universidade

privada, os dados de estudantes atendidos foram incluídos, independentemente de terem recebido ou não diagnóstico pelo DSM.

Uma amostra do *Employee Assistance Program* (EAP), gerenciado pelo Human Affairs International (HAI), foi composta por pacientes, que procuraram ou foram encaminhados para a assistência e receberam um diagnóstico do DSM-III-R. Os pacientes, que buscaram ajuda ou foram encaminhados, mas não foram diagnosticados ou tratados por um problema emocional, foram excluídos da amostra. Também foram excluídos os pacientes que foram imediatamente encaminhados para internação ou tratamento ambulatorial. Os dados referem-se a pacientes de sete estados americanos.

A amostra de pacientes em tratamento domiciliar foi proveniente de uma Clínica de Cuidados Domiciliares pertencente a uma universidade, usada no treinamento profissional de clínicos em serviço social, de psicólogos clínicos, de terapia familiar e de casal. Também foram incluídos pacientes inseridos em um programa comunitário de um centro saúde mental na área rural do estado de Ohio.

A amostra de pacientes internados para tratamento clínico foi composta por informações oriundas de Utah e Massachusetts. Os dados desses pacientes foram combinados em uma única amostra, já que os valores dos grupos eram comparáveis.

A análise da confiabilidade pelo método de teste-reteste com a amostra de 157 estudantes universitários de uma grande universidade americana. A média de idade da amostra foi 23,04 (DP = 3,41), dos quais 34,3% eram homens e 65,7% mulheres. Os resultados mostraram que o OQ-45 tem boa estabilidade temporal, paras o domínio SD (r = 0,78), IR (r = 0,80), SR (r = 0,82) e escore total (r = 0,84). Segundo Lambert et al. (1996), os resultados do reteste mostraram uma estabilidade temporal do instrumento, tanto para o escore total da escala como para as três subescalas.

Também foi realizada a análise da consistência interna com a amostra de estudantes e com os dados de uma amostra de 298 pacientes da EAP empregando o alfa de Cronbach. Os coeficientes obtidos com a amostra de estudantes de graduação foram: SD = 0,92; IR = 0,74; SR = 0,70 e OQ Total = 0,93. Na amostra de pacientes os dados são semelhantes SD = 0,91; IR = 0,74; SR = 0,71 e OQ Total = 0,93. Tanto a pontuação total como a dimensão Desconforto Subjetivo mostraram excelente consistência interna (0,93 para total e de 0,91 a 0,92 para SD). Entretanto maior heterogeneidade foi encontrada no conteúdo das subescalas Relações Interpessoais e desempenho do Papel Social de (0,70 a 0,74), o que não é surpreendente, dada a amplitude do funcionamento que estas subescalas tentam avaliar. Quanto à precisão relativa ao reteste semanal dos estudantes os coeficientes no reteste

variaram de 0,66 a 0,86, diminuindo progressivamente, conforme passava o tempo.

Foi estudada a validade concorrente pela correlação dos escores do OQ-45.2 com outras medidas que avaliam constructos similares, indicadas a seguir: *Symptom Checklist-90-Revised* (SCL-90), Inventário de Beck de Depressão (BDI), *Zung Self Rating Depression Scale* (ZSDS), *Taylor Manifest Anxiety Scale* (TMAS), *State*-Trait Anxiety Inventory (STAI), (STAI-Y1= State Anxiety), (Y-2 = Trait Anxiety), *Inventory of Interpersonal Problems* (IIP), *Friedman Well-Being Scale (FW-B)* e *Social Adjusment Scale Self-Reported* (SAS-SR).

Foram obtidas seguintes correlações: para a dimensão SD com o escore geral do SCL-90R = 0,76; BDI =0,63; ZSDS = 0,88; ZSAS = 0,81; TMA = 0,88; STAI-Y1 = 0,50; STAI-Y2 = 0,65; IIP = 0,64. Para a dimensão IR com: SCL-90R = 0,53; IIP = 0,62 e FW-B = 0,48. Para a dimensão SR com o SCL-90R = 0,47; IIP = 0,51 e SAS = 0,43.

Para o escore total com o SCL-90R = 0,73; BDI = 0,80; ZSDS = 0,88; ZSAS = 0,81; TMA = 0,86; STAI-Y1 = 0,64; STAI-Y2 = 0,80; IIP = 0,66; SAS =0,65; SF-36 = 0,81 e FW-B = 0,81.

Para a validade de constructo o autor investigou a capacidade do instrumento para discriminar entre os níveis de psicopatologia, sua sensibilidade às mudanças, a estrutura de fatores e a relação das subescalas com a pontuação total. Os resultados mostraram uma explícita diferença entre os valores dos escores médios das amostras de pacientes, que foram significativamente mais altas do que de não pacientes (Umphress, Lambert, Smart, Barlow & Clouse, 1997). Não foram encontradas diferenças significativas relativas ao gênero, nas duas amostras.

Estudos Internacionais

Várias pesquisas foram realizadas com a finalidade de avaliar as propriedades psicométricas do OQ-45, na Alemanhã, Chile, Itália e Holanda. Doerfler, Addis e Moran (2002) investigaram a validade divergente e concorrente e a sensibilidade para mudanças, em pacientes psiquiátricos adultos. Utilizaram a *Symptom Identification Scale* (BASIS-32) e o *Outcome Questionnaire* (OQ-45). Na comparação das medidas observou-se que os dois instrumentos são sensíveis para avaliar mudanças, ambos mostraram evidências de validade convergente e divergente nas subescalas específicas, os escores totais entre os dois testes mostraram alta correlação. As evidências de validade concorrente foram obtidas com outras medidas utilizadas como critério, os resultados sugerem diferenças entre os grupos diagnósticos nas subescalas específicas.

A escala de avaliação de resultados OQ-45.2 foi adaptada em outros países. No Chile,

os estudos mostraram evidências de que ela é útil na mensuração de resultados transculturais. De La Parra e Bergen (2002) relataram o sucesso da sua aplicação em pacientes de classes de renda média e baixa que receberam tratamentos nos casos de internação ou em regime ambulatorial. Observaram semelhanças entre as amostras de universitários dos Estados Unidos (Média= 42,33; Desvio Padrão= 16,60, para N= 238) e os dados obtidos com a população chilena (Média = 48,7; e Desvio Padrão = 19,3, para N= 129). Algumas pequenas alterações foram necessárias para traduzir o OQ-45.2 para o espanhol falado no Chile, a fim de adequar o instrumento ao nível de educação particularmente baixo de alguns dos participantes e diferenças na estrutura das frases no idioma espanhol. Os autores constataram que o coeficiente de confiabilidade e a validade foram bastante elevados, semelhantes aos encontrados nos Estados Unidos. Foram obtidos coeficientes alfa de consistência interna (0,91) para os dois grupos, pacientes (N= 124) e controle (N= 202) e confiabilidade pelo método teste-reteste, a segunda aplicação ocorreu após um intervalo de duas semanas, para SD (0,90), IR (0,79) e SR (0,82) e OQ Total (0,90), com valores quase idênticos aos obtido no estudo americano e publicados no manual do OQ-45.2. Além disso, a versão em espanhol mostrou sensibilidade aos efeitos do tratamento, que evidenciaram grandes mudanças nos pacientes que completaram o tratamento, mudanças moderadas nos que estavam em terapia e pouca ou nenhuma mudança nos que não foram submetidos a tratamento.

Jong et al. (2007) fizeram uma validação transcultural do OQ-45 na Holanda, na qual verificaram as propriedades psicométricas e de equivalência entre a estrutura dos fatores da adaptação da versão holandesa com a versão original americana. A amostra foi composta de universitários (N=268), pessoas da comunidade (N=810) e de três organizações de saúde mental (N=1920). Os resultados mostraram que as propriedades psicométricas da versão holandesa foram adequadas e semelhantes às do instrumento original. No entanto foram encontradas algumas diferenças relativas à equivalência. Na análise fatorial foram obtidos dois fatores adicionais, um composto por itens de papéis sociais e outro que reflete a ansiedade e sintomas somáticos. Além disso, houve diferenças entre os dados normativos das amostras holandesa e americana, o que resultou em uma nota de corte diferente para estimar alterações clinicamente significativas nos dados holandeses.

Lo Coco et al. (2008) traduziram o OQ-45.2 para a Itália e o aplicaram em duas amostras. A primeira, composta por estudantes (N=522) com idade média de 26,06 e desvio padrão 9,66, sendo 65,3% mulheres e 34,7% homens. A segunda foi de pacientes psiquiátricos (N=301), dos quais 64,5% de mulheres e 35,5% homens, com idade média de 33,89 e desvio padrão de 14,61. Todos os pacientes preencheram os critérios diagnósticos

estabelecidos no DSM-IV. Os transtornos de humor foram os mais frequentes (43,5%), seguidos de transtornos de ansiedade (17,6%), transtornos alimentares (7,3%) e transtorno de personalidade (6%). O OQ-45 foi aplicado nos pacientes antes da primeira sessão de psicoterapia. O estudo das propriedades psicométricas da versão em italiano da OQ-45 indicaram que elas eram adequadas e semelhantes às do instrumento original. A consistência interna pelo alfa de Cronbach foi de 0,92. A validade concorrente dos escores do OQ-45 estabelecidos pelas correlações com BDI (0,75), com STAY Y1 (0,69) e STAY Y2 (0,81) foram satisfatórias. As correlações entre os itens das subescalas mostraram uma boa consistência interna pelo alfa de Cronbach OQ total (0,93), SD (0,92), IR (0,77) e SR (0,66).

Os Estados Unidos apresentam uma extensa diversidade multicultural com diferentes etnias, o que exige esforços para desenvolver instrumentos para avaliar a personalidade e a inteligência, uma vez que que essas diferenças, podem dificultar a tomada de decisões clínicas, como diagnóstico e planejamento da psicoterapia. Muitos testes foram padronizados para pacientes europeus e americanos, sendo de fundamental importância novos estudos para validar medidas para diferentes contextos e etnias. Preocupados com o excesso de patologização de pacientes oriundos de culturas diversas, Beretvas, Kearney e Barón (2003) estudaram as evidências de validade e confiabilidade da pontuação no OQ-45 com diferentes amostras. Foi investigada a estrutura fatorial em grupos de brancos e não brancos. Na análise fatorial confirmatória os itens foram agrupados em quatro fatores que incluíram depressão/ansiedade, sentimentos de bem-estar, impacto do estresse e abuso de álcool. As correlações entre os quatro fatores variarem de 0,223 a 0,783 ($p \leq 0{,}001$). A correlação mais forte foi entre o fator Depressão/ansiedade e os sentimentos de bem-estar ($r = 0{,}783$). Além disso, como esperado, houve correlações moderadamente fortes entre o fator impacto de estresse e do fator de depressão / ansiedade ($r = 0{,}697$) e com sentimentos de bem-estar ($r = 0{,}551$). As correlações entre o fator Abuso de álcool e os outros fatores foram baixos (variando de 0,223 a 0,260). Os autores ressaltaram a necessidade de novos estudos para avaliar a confiabilidade das subescalas.

O OQ-45.2 também é usado como medida para validar outros instrumentos, como no estudo de Draper, Jennings, Barón (2003) estudo da validade concorrente do *Presenting Problems Checklist*, que foi aplicado em um consórcio de pesquisa envolvendo mais de 50 centros de aconselhamento dos Estados Unidos. Os autores pretendiam avaliar se a estrutura fatorial do *Cheklist* é adequada para fins clínicos e de pesquisa, e a validade do instrumento. Empregaram a Análise Fatorial Exploratória (AFE) para extração dos fatores latentes que explicassem a correlação entre os itens. As variáveis que compõem cada fator passaram por uma

nova análise fatorial com rotação ortogonal. Após o agrupamento dos itens do *Checklist*, cada um dos cinco fatores foi correlacionado com a pontuação total da primeira aplicação do OQ-45.2 com a da última sessão. Os resultados mostraram uma excelente consistência interna para o *Checklist* e a diferença no coeficiente de correlação entre os dois instrumentos foi estatisticamente significativa, indicando evidências para a validade concomitante do *Checklist* com o nível de desconforto geral medido pelo OQ-45, o que mostrou a validade concorrente, isto é, os dois instrumentos estão medindo sintomatologias semelhantes.

Foram realizados vários estudos para avaliar os ganhos na psicoterapia como o de Kearney, Draper e Barón (2003), que avaliaram os benefícios obtidos por minorias étnicas nos campi universitários. A amostra foi composta de 1.166 estudantes afro-americanos, asiático-americanos, caucasianos, hispânicos, que buscaram ajuda nas mais de 40 universidades nos Estados Unidos. Os estudantes asiáticos apresentaram maior angústia na avaliação inicial, seguidos por hispânicos, africanos e estudantes caucasianos. Todos os grupos pareceram beneficiar-se da psicoterapia, entretanto nenhum deles preencheu os critérios de mudanças clinicamente significativas.

Segundo Haug, Puschner, Lambert e Kordy (2004), a monitoração contínua de resultados de psicoterapia requer instrumentos curtos, válidos e com alta sensibilidade para avaliar mudanças. Eles investigaram a adequação da versão alemã do OQ-45.2, o *Ergebnistfragebogen* (EB-45). A amostra foi composta por pessoas da comunidade (N = 232), passageiros ferroviários (N = 615), pacientes de um ambulatório de psicoterapia (N = 671) e pacientes internados (N =960). Para avaliar a validade concorrente do EB-45, sua capacidade de discriminar amostras de paciente e não pacientes, sua sensibilidade para avaliar mudanças e a sensibilidade e especificidade utilizaram como critério externo pensamentos suicidas e a capacidade de trabalho. Encontraram uma concordância moderada (0,40) com o *Symptom Checklist* e o *Brief Symptom Inventory*. O EB apresentou boa capacidade para discriminar entre pacientes e não pacientes, confirmando a sua adequação para monitoramento contínuo das mudanças em psicoterapia.

Reese, Norsworthy e Rowlands (2009) estudaram o uso de *feedback* para monitorar o progresso do tratamento e melhorar o resultado da psicoterapia. Seu objetivo era verificar se o uso de *feedback* (medidas para avaliação de resultados e aliança terapêutica) contínuo promove mudança confiável. Os pacientes respondiam o OQ-45.2 em um terminal de computador, logo que chegavam à clínica, e era gerado um relatório enviado ao terapeuta. Os dados obtidos foram examinados pelo terapeuta e paciente durante as sessões. As amostras foram compostas de pacientes que passaram por terapia individual em um centro universitário

(N= 74) ou em uma clínica de treinamento de pós-graduação (N= 74). Os resultados indicaram que 50 % dos pacientes que receberam tratamento com terapeutas em que ambos discutiam o progresso da terapia demonstraram mudança confiável após a 7ª sessão ou 9ª sessão para os pacientes atendidos no centro de aconselhamento universitário. As médias foram comparadas por análise de variância, indicando mudanças estatisticamente significativas. O sistema de *feedback* foi desenvolvido por Miller & Duncan (2004) e denominado de *Partners for Change Outcome Management System* (*PCOMS*), os autores argumentam que a vantagem no procedimento em que o terapeuta e o paciente examinam dados das avaliações, especialmente no início da sessão, levam a uma resposta imediata na própria sessão.

O estudo de Lara, Cruz, Vacarezza, Florenzano e Trapp (2008) comparou os resultados obtidos com 25 pacientes internados em um Hospital Psiquiátrico com 25 pacientes de um hospital dia. Os instrumentos OQ-45 e InterRAI-Saúde Mental foram aplicados quando os pacientes deram entrada nos hospitais. Os resultados mostraram que 29 itens de ambos os instrumentos se referem a variáveis semelhantes, dos quais oito apresentaram correlações estatisticamente significativas, com os itens do OQ45 relacionados à sintomatologia depressiva e ansiosa. Não houve associação entre o os dois instrumentos nos itens que se referem ao papel social (SR) ou relações interpessoais (RI). Os autores destacam a necessidade de estudos com outras amostras.

Em seu trabalho Okiishi (2006) analisou os dados coletados em uma amostra de 5.000 clientes atendidos por 71 terapeutas durante um período de seis anos em um Centro de Aconselhamento Universitário. O OQ-45.2 foi aplicado em cada sessão para acompanhar as reações dos clientes ao tratamento. Foram também coletadas informações referentes à orientação teórica dos terapeutas, gênero, tempo de experiência e tipo de treinamento. Os dados foram analisados a partir da modelagem hierárquica linear (HLM) para determinar se os traços gerais do terapeuta (ou seja, orientação teórica, tipo de treinamento) foram responsáveis pelas diferenças na melhora dos clientes. A HLM é uma metodologia de análise de dados que permite explicar a variabilidade da variável resposta pelas variáveis preditoras incluídas em diferentes níveis hierárquicos e que possa ser quantificável em cada nível, tal que a proporção da variabilidade possa ser comparada diretamente. Os resultados desta análise mostraram que o tipo de treinamento (aconselhamento psicológico, psicologia clínica, terapia familiar, casal), os anos de formação, a orientação teórica (comportamental, cognitivo-comportamental, humanista, psicodinâmica), e o sexo não contribuem significativamente na diminuição do tempo para os clientes melhorarem. Entretanto, um olhar

mais atento aos terapeutas mais eficientes pode ser possível em pesquisas futuras, desde que estes profissionais estejam dispostos a identificar-se e serem estudados.

Potter, Baker, Sanders e Peterson (2009) destacam que poucos estudos avaliaram a eficácia de tratamentos de saúde mental quanto às reações de estresse em militares em seus locais de trabalho. Realizaram uma investigação piloto para avaliar a eficácia de um tratamento de controle de estresse, com 38 militares (33 homens e cinco mulheres) que passaram pelo programa conduzido na CSC do Exército, na Base Aérea de Balad (Iraque). O programa incluiu psicoterapia individual e de grupo para redução dos sintomas de estresse relativos aos combates operacionais e para possibilitar o desenvolvimento de estratégias de enfrentamento, habilidades para gerenciamento de estresse e relações interpessoais. Foram utilizados os seguintes instrumentos: o OQ-45.2 e a Escala de Transtorno de Estresse Pós-traumático *(Posttraumatic Stress Disorder Checklist-Military - PCL-M),* desenvolvida por Weathers et al. (1993). Os resultados mostraram reduções significativas nos escores obtidos, nos dois instrumentos (OQ-45.2 e Escala de Estresse Pós-Traumático), indicando os benefícios do tratamento para militares expostos a situações estressoras.

Segundo Florenzano, Artero, Gonzalez, Martinez & Ortega (2005), os estudos para avaliação de resultados de psicoterapia são importantes para auxiliar a tomada de decisões no setor de saúde. Para demonstrar que o tratamento é benéfico para o paciente, é necessário medir a dimensão e o tipo de mudanças que ocorreram na psicoterapia. É essencial o uso de instrumentos padronizados que sejam eficazes na avaliação de mudança terapêutica, por meio de indicadores observáveis e mensuráveis das diferentes áreas da vida do indivíduo. O OQ45 foi utilizado para avaliar a eficácia do processo terapêutico e resultados quantitativos do sucesso ou fracasso das intervenções, em um hospital psiquiátrico. Foram avaliados 319 pacientes atendidos em triagem no Serviço de Psiquiatria do Hospital de Salvador (SPHDS), além de 40 pacientes na primeira, segunda e terceira sessões, 257 casos de pacientes hospitalizados no SPHDS, 191 pacientes atendidos em um serviço de Psicologia de uma universidade chilena. Todos os pacientes foram avaliados em três momentos: no momento de iniciar o tratamento, três meses depois de ter começado e na alta ou seis meses após o início tratamento. Nos grupos estudados confirma-se que a escala OQ-45.2 foi um instrumento sensível para avaliar a mudança terapêutica a curto prazo.

Segundo Beckstead et al. (2003), o método de Jacobson é usado frequentemente em estudos para avaliar mudanças consideradas clinicamente significativa, mas não tem pesquisas sobre a validade do *Reliable Change Index* (RCI) de Jacobson e Truax. Os autores referem-se a dois estudos que examinaram a validade de mudança utilizando o RCI. O valor

do RCI é obtido por subtração da pontuação obtida antes do início do tratamento em relação à obtida no pós-tratamento, dividindo pelo erro padrão de medida. O primeiro estudo, de Ankuta e Ables (1993) comparou os pacientes, que pelo método de Jacobson apresentaram mudança clinicamente significativa, que se mostraram mais satisfeitos com o tratamento com os que não preencheram os critérios para mudança. O segundo, de Lunnen e Ogles (1998) ampliou o estudo de Ankuta e Ables, realizando uma análise multivariada dos componentes do método de Jacobson e Truax (RCI). Eles dividiram pacientes ambulatoriais em três grupos com base na mudança: os que tiveram ganhos, o que não apresentaram mudanças e os que pioraram com o tratamento. Os pacientes dos três grupos classificaram sua mudança percebida, satisfação com o tratamento e aliança terapêutica. Os dados obtidos indicaram que a mudança percebida e a aliança terapêutica foram significativamente maiores para o grupo que obteve ganhos. A satisfação com os serviços não diferiu entre os grupos. Os instrumentos de medida não distinguiram os grupos sem mudanças dos que pioraram. Os resultados mostraram que o RCI é um método eficaz de avaliação de melhora sintomática, mas um indicador menos válido para os casos em que o paciente obtém resultados negativos com a psicoterapia. Ogles et al. (2001) destacam a necessidade de mais pesquisas sobre a validade do método de Jacobson, como meio de operacionalizar o conceito de mudança significativa.

Em resposta a essa necessidade de mais pesquisas sobre o conceito de mudança clinicamente significativa, Beckstead et al. (2003) investigaram a correspondência entre como diferentes medidas de resultado classificaram os pacientes, antes e após o tratamento, em dois grupos como funcional e disfuncional de acordo com a pontuação obtida e que permitem detectar dificuldades em áreas específicas. Os pacientes que obtiveram classificação indicando mudança confiável também foram examinados. A classificação de mudança obtida pelos pacientes como confiável também foi examinada. Foram utilizados os seguintes instrumentos: *Symptom Checklist-90-R, Social Adjustment Rating Scale-Self Reported/e Other Report, Inventory of Interpersonal Problems-Short Form, O Quality of Life Inventory* e *Client Satisfaction Questionnaire-8*. Os autores partiram da hipótese de que haveria concordância significante entre a classificação de cada paciente como funcional ou disfuncional com base na nota de corte do OQ-45.2 e as notas de corte derivadas de cada uma das quatro medidas. A segunda hipótese foi de que a classificação de cada cliente em relação aos critérios de Jacobson para mudanças como clinicamente significativas seriam concordantes. Finalmente, a última considerou que os clientes que tiveram mudanças ficaram mais satisfeitos com o tratamento do que os que pioraram.

Os resultados da pesquisa de Beckstead et al. (2003) forneceram evidências para a validade da construção do conceito de significância clínica e a nota de corte do OQ-45.2 foi de 45 pontos. Houve concordância entre a classificação dos pacientes como tendo alcançado mudança clinicamente significativa a partir da pontuação no OQ-45.2, com a classificação obtida a partir dos resultados de outros instrumentos que indicaram mudança clinicamente significativa. O conceito de mudança clinicamente significativa é importante para diminuir a distância entre a pesquisa e a prática clínica, possibilitando uma atenção aos resultados individuais do cliente ao invés de compará-lo com as médias de grupo.

A psicoterapia é objeto de estudo em diferentes contextos, segundo Lambert et al. (2006), nos últimos anos, a atenção centrou-se na prestação de serviços de psicoterapia para minorias étnicas. Como a população de minorias étnicas nos Estados Unidos continua a crescer, maiores exigências serão colocadas sobre as agências de saúde mental para fornecer serviços adequados. O censo nos EUA estima que aproximadamente 30% da população é composta por minorias étnicas e, na Califórnia, elas constituem agora a maioria. As evidências sugerem que as minorias étnicas enfrentam uma maior quantidade de eventos estressores por causa do nível socioeconômico mais baixo, falta de moradia, desemprego, pressões para a aculturação e situações de preconceito e discriminação. Um aspecto de interesse no contexto na psicoterapia é o papel da semelhança étnica entre o cliente e o terapeuta. Alguns têm defendido essa correspondência como um meio de melhorar os serviços, como Sue, Fujino, Hu, Takeuchi & Zane (apud Lambert, 2006). Um dos estudos mais abrangentes de meta-análise, sobre a correspondência étnica entre terapeutas e pacientes, encontrou tamanhos de efeito insignificantes associados a taxas de abandono do tratamento (r = 0,03), assiduidade no atendimento (0,04) ou nível de funcionamento (r = 0,01), entre os pacientes afro-americanos, hispânicos, asiáticos americanos, brancos que foram pareados ou não com terapeutas de origem étnica semelhante.

O estudo de Lambert et al. (2006) foi realizado como pesquisa-ação com o objetivo de decidir se um Centro de Aconselhamento da Universidade deveria gastar tempo e recursos adicionais para treinamento de terapeutas em sensibilidade multicultural, a fim de melhorar o resultado obtidos pelas minorias étnicas em psicoterapia. Esta foi uma questão especialmente importante para o Centro por causa da escassez tanto de pacientes e funcionários pertencentes às minoritárias étnicas. A pergunta de pesquisa foi: Os clientes de minorias étnicas e atendidos no centro de aconselhamento têm os resultados do tratamento diferentes dos clientes brancos? Os estudos de correspondência étnica geralmente dependem da declaração dos clientes de pertencerem a uma minoria étnica e restringir identidade com as seguintes

categorias: (1) Africano americano ou preto, (2), da Ásia ou asiático-americano, (3) Nativo Americano ou Índio Americano, (4) Hispânico ou latino americano, e (5) Europeu americano Caucasiano ou branco. Zane, Hall, Sue, Young e Nunez (2004) notaram a escassez de estudos com clientes índios americanos e nativos do Alasca e/ou terapeutas pertencentes a uma das etnias entre os grupos citados anteriormente. Lambert et al. (2006) estudaram os resultados de tratamentos de clientes americanos nativos. Foi obtida uma de linha de base com os resultados de clientes de diferentes etnias para avaliar os efeitos da intervenção. Os dados arquivados foram analisados para avaliar os resultados do tratamento em quatro grupos étnicos minoritários. Os clientes que se identificaram como nativos americanos, latino, Africano, ou ilhéu da Ásia/Pacífico e que tinham pré e pós-teste com os resultados do OQ-45.2 foram selecionados para o estudo. Cada membro do grupo étnico foi combinado com um cliente identificado como Europeu americano ou branco (grupo controle), com base no nível inicial de perturbação, gênero, estado civil e idade. Foram realizadas análises de variância univariadas para examinar as diferenças entre os pré e pós-testes da amostra total. Além disso, quatro grupos foram separados e calculados os testes *t* para avaliar as diferenças de resultados entre o grupo de comparação étnica com o grupo controle (caucasiano). Esses resultados foram interpretados indicando que os clientes com receio de sua participação em um grupo étnico, melhoraram significativamente do início ao fim do tratamento. Entretanto isto é típico dos clientes que passam por tratamento breve. As diferenças entre cada um dos oito grupos de tratamento quanto aos seus ganhos com os tratamentos não atingiram significância estatística ($F = 0,69$, $df = 7$, $p = 0,68$). No entanto, algumas diferenças entre os grupos chamaram a atenção. O tamanho dos efeitos no pré e pós-teste variou de um mínimo $d= 0,33$ em Latino, a $d = 0,61$ em Nativos Americanos. A redescoberta do uso da magnitude do efeito nos estudos em Psicologia representa uma tentativa de substituir o conceito de significância estatística com noções mais úteis de significância prática específica.

Os pesquisadores Zane et al. (2004) destacaram os achados relativos aos clientes Nativos Americanos que foram atendidos em várias sessões. Surpreendentemente esses clientes (N=50) tiveram os resultados mais positivos. Os autores ressaltaram a importância de levantar hipóteses para compreender as razões dos seus resultados.

As situações em que os tratamentos oferecidos por terapeutas, que em sua maioria não apresentam correspondência étnica com o paciente, podem trazer possíveis consequências negativas para clientes pertencentes a um determinado grupo étnico. Uma estratégia para lidar com estes casos é oferecer uma formação com vista a melhorar a competência multicultural pela educação multicultural. Os autores concluíram que é necessário avaliar a

necessidade de educação multicultural e fornecer uma estimativa inicial da eficácia para a comparação dos resultados dos esforços futuros para a melhora no tratamento das minorias étnicas.

A pesquisa de Ballesteros, Vitriol, Florenzano, Vacarezza e Calderón (2007) objetivou analisar a relação entre trauma na infância e depressão. Eles trabalharam com uma amostra de 136 pacientes, dos quais 87 tinham um histórico de politrauma infantil e 49 sem eles. Todos foram avaliados na admissão com a Escala de Depressão de Hamilton (HAM-D), Escala de Trauma de Marshall (número de respostas afirmativas para eventos traumáticos) e OQ-45.2. Compararam a gravidade dos sintomas depressivos, área de relações interpessoais e papel social em pacientes com depressão grave, de acordo com a presença e número de eventos traumáticos. Os resultados mostraram que as médias no OQ Total foram 113,7, para pacientes com mais de três eventos traumáticos e 85,6 em pacientes com dois eventos ou menos. As diferenças foram estatisticamente significativas ($t = 5{,}329$, $p \leq 0{,}001$). Os pacientes com depressão grave, nos subgrupos de pacientes com trauma de infância, mostraram maior associação entre a gravidade dos sintomas depressivos nos casos com três ou mais antecedentes de politraumas na infância e abuso infantil. Os escores foram maiores nas subescalas e no OQ Total, quando comparadas com o grupo de mulheres com depressão severa, mas com menos eventos traumáticos na infância.

De acordo com Dell´Aquila, Castro e Mosca (2009), a obesidade no Chile tem se transformado nos últimos 10 anos em um problema de saúde pública, afetando adultos, crianças e adolescentes. A obesidade está associada a fatores socioculturais, como vida sedentária e falta de espaço físico para realização de atividades de lazer. Os estudos assinalam a presença de fatores emocionais, mas não trazem maiores referências a respeito, assim os autores buscaram identificar os fatores emocionais em adolescentes com obesidade e, após o psicodiagnóstico, encaminhar para psicoterapia. Para avaliar a vida emocional e relacionamentos significativos desses adolescentes foram aplicados os seguintes instrumentos para o diagnóstico: entrevista clínica, Teste das Relações Objetais (TRO), Teste de Apego ou Instrumento de Vínculo Parental (PBI) e OQ-45.2, nos dois grupos de adolescentes com e sem diagnóstico de obesidade (grupo controle), 25 em cada grupo, com uma amostra de 25 jovens entre 14 e 18 anos em cada grupo. O grupo de adolescentes com obesidade apresentou maiores índices de ansiedade e depressão e maiores dificuldades em suas relações familiares e sociais. O sobrepeso foi associado com acontecimentos importantes na vida dos adolescentes, tais como morte de pais e avós, acidentes e doenças com pessoas da família, mudanças de cidade e escola, e presença de sintomatologia psíquica obtidos com o OQ-45.2.

Além das pesquisas para monitorar o progresso do tratamento, a avaliação contínua tem sido estendida para investigar como maximizar o uso dos resultados na psicoterapia. Quando os resultados das avaliações ao longo do processo são compartilhados entre o terapeuta e o paciente em cada sessão, a psicoterapia é mais eficiente, do que nas situações em que apenas com os terapeutas recebem essas informações. As medidas de aliança terapêutica, estágios de mudança e apoio social aumentam a efetividade do tratamento, indicando que o uso contínuo dos dados de resultados é benéfico para todos os clientes.

Os pesquisadores buscaram identificar qual o número de sessões suficiente para provocar mudanças nos pacientes, sendo este tipo de estudo chamado de dose-efeito. As relações dose-efeito em Psicoterapia Breve foram avaliadas com base em estudos de centros de aconselhamento universitário. Vários exemplos nos anos 1950 e 60 foram examinados para verificar a relação entre o número de sessões de clientes atendidos e os ganhos obtidos. Os resultados indicaram que quanto maior o número de sessões, maior o ganho (Johnson, 1965, apud Hansen, Lambert & Forman, 2002). Estudos mais recentes têm confirmado estes achados iniciais, concluindo que a duração do tratamento é positiva e associada à melhora e benefício terapêutico (Kadera, Lambert & Andrews, 1996).

O uso de medidas de resultados de Pesquisa em Psicoterapia vem sendo objeto de interesse dos profissionais que atuam em centros de treinamento de psicoterapeutas. A pesquisa de Callahan e Hynan (2005) objetivou identificar, se o modelo dose-efeito é aplicável na clínica de treinamento de terapeutas. O estudo foi realizado em um programa de doutorado em Psicologia Clínica para formação de terapeutas, que testou este modelo, isto é, a relação entre o número de sessões realizadas e o efeito terapêutico. Este modelo de psicoterapia surgiu a partir de estudos de meta-análise, realizados por Howard, Kopta, Krause e Orlinsky (1986), abrangendo 30 anos de pesquisas e tem gerado um interesse renovado nas pesquisas em psicoterapia. Entretanto, os autores não examinaram possíveis diferenças relacionadas ao contexto onde os atendimentos foram realizados, por exemplo, se o local de atendimento foi um centro de formação. Os dados foram obtidos com o OQ-45.2 aplicado antes de cada sessão durante todo o tratamento (número total de sessões em média 14,81 sessões). A amostra foi composta por 61 pacientes atendidos por 21 estagiários. Apesar dos resultados relatados Kadera, Lambert e Andrews (1996) de que como o indivíduo em formação era presumivelmente menos experiente, os resultados mostraram o surgimento de padrão similar aos outros estudos, entretanto a resposta dos pacientes ao tratamento não foi tão rápida quanto o modelo de dose-efeito poderia prever.

Segundo Kadera et al. (1996) um dos primeiros estudos que despertou o interesse de

pesquisadores na relação entre a duração do tratamento e benefício terapêutico foi a de Seeman, em 1954. Ele trabalhou com uma pequena amostra de pacientes (estudantes universitários) e concluiu que o ganho terapêutico, com base na avaliação dos terapeutas, foi o maior para pacientes que permanecem em tratamento por pelo menos 20 sessões. Este trabalho foi seguido por vários estudos com pequenas variações na metodologia e produzindo resultados semelhantes. As investigações demonstraram que os ganhos obtidos na psicoterapia estavam relacionados à duração da psicoterapia, o que é compatível com a literatura atual. Entretanto, a metodologia utilizada na época é inadequada para os padrões atuais, porque a dependência exclusiva de amostras de pacientes provenientes centros de aconselhamento universitário limita a generalização dos resultados. Atualmente o ressurgimento do interesse na relação entre o benefício e duração da terapia foi estimulado pela constante referência ao estudo de Seeman (apud Howard, Kopta, Krause & Orlinsky, 1986). Esses pesquisadores propuseram uma psicoterapia "dosagem" como modelo, em que a dose foi definida como o número de sessões, e o efeito foi definido como a porcentagem de pessoas que se espera que tenham melhora em cada sessão. O número Dose-efeito em psicoterapia foi obtido pelos resultados de 15 estudos sobre avaliação de resultados de psicoterapia retirados da literatura. Os resultados indicaram que em aproximadamente l5% dos pacientes poderia ser esperada na melhora antes da primeira sessão (presumivelmente através de remissão espontânea), 25% após a 1ª sessão, 50% depois oito sessões, e cerca de 75% após 26 sessões. A análise indicou ainda que os pacientes deprimidos tiveram as mais baixas doses de tratamento, seguidos pelos com neuroses de ansiedade e os *borderline* que necessitaram de mais sessões.

No estudo de Kadera et al. (1996) com uma amostra de 45 clientes, usando a pontuação total do OQ-45.2, os autores relataram que 2% dos pacientes preencheram os critérios de acordo com RCI de Jacobson para recuperar-se na 2ª Sessão, 18% na 7ª sessão, e 46% foram recuperados em 26 sessões. Enquanto na pesquisa de Callahan e Hynan, 2% dos pacientes mostraram recuperação na 2ª sessão, 3% na 7ª sessão e 16% em 26 sessões. Os autores destacaram as limitações da investigação, no entanto, ressaltando que são limitações presentes na literatura. Essas limitações incluem o registro de um diagnóstico primário para cada cliente, do uso ou não de medicamentos psicotrópicos, do papel do planejamento da psicoterapia e do custo do tratamento.

A literatura não apresenta informações sobre a eficácia de tratamento de pacientes com problemas auditivos, com ou sem intérprete. De acordo com Storch (2010) a acessibilidade aos serviços de saúde mental e tratamento é ineficaz para as pessoas com perda

auditiva total, principalmente por causa das barreiras de comunicação entre o terapeuta e o paciente. Em seu estudo a autora investigou a eficácia da psicoterapia para pacientes surdos, comparando os resultados dos que trabalharam com um terapeuta que utilizava linguagem de sinais, com outros que trabalharam com intérpretes. O OQ-45.2 foi aplicado para avaliar o nível funcional do participante, do processo de mudança, do desconforto subjetivo, do funcionamento interpessoal e do desempenho dos papéis sociais em uma amostra de conveniência de 13 pacientes surdos com idades entre 16 e 60 anos. Os resultados indicaram uma melhora dos pacientes atendidos pelo terapeuta que utilizou sinais como forma de comunicação, mas os pacientes surdos atendidos por terapeuta que usou intérprete não tiveram melhora. O autor enfatiza a necessidade de uma melhor compreensão da cultura dos pacientes portador de deficiência auditiva, de uma compreensão mais profunda sobre a cultura "surda". Storch (2010) aponta como limitação do estudo o uso de um instrumento que não corresponde ao nível gramatical na qual o paciente "surdo" está inserido. Destaca ainda, a questão do interprete não ter do treinamento para atuar em sessões de psicoterapia.

O número de pacientes que abandonam ou interrompem o atendimento é objeto de preocupação em diferentes países, enquanto no Brasil são encontrados vários estudos, caracterizando a população atendida em clínica-escola (Chilelli & Enéas, 2000; Gatti & Jonas, 2007; Campezatto & Nunes, 2007; Benetti & Cunha, 2008; Bortolini, Pureza, Andretta & Oliveira, 2011).

Tanto nos artigos de periódicos como nas teses e dissertações na área de psicoterapia observa-se um interesse nos pacientes que não se beneficiam com o tratamento, cerca de 5 a 10% dos pacientes (Lambert & Ogles, 2004), como naqueles que desistem do atendimento, em torno de 30 a 60 %. Estes casos precisam ser compreendidos por meio de uma integração entre a pesquisa e a prática clínica, direcionando o foco para os resultados do paciente individualmente e não de grupos. A proposta atual enfatiza a importância do monitoramento de resultados com a finalidade de planejamento do tratamento e da innclusão de indicadores que permitam identificar e atender melhor os pacientes com risco de resultado negativo.

No estudo de Shimokawa (2010) ele refere-se a uma minoria da população de pacientes que não se beneficiam com o tratamento. Ele foi buscar dados que possam possibilitar a melhora dos resultados em pacientes com risco de piorar com o tratamento. Foram analisados tratamentos por meio de uma meta-análise de um conjunto de dados agrupados oriundos de seis grandes estudos realizados em um grande centro universitário e ambulatório hospitalar (N = 6151) para verificar os efeitos do uso de *feedback* no progresso dos pacientes. Foram utilizados os escores do OQ-45.2 para monitorar o progresso

terapêutico e para identificar precocemente os pacientes com risco de resultados negativos na psicoterapia. Os profissionais responsáveis pelo atendimento recebiam informações para que pudessem intervir antes de ocorrer fracasso no atendimento (efeito negativo). As análises da eficácia das intervenções de *feedback* tinham o objetivo de obter as estimativas dos efeitos previsíveis pela implementação deste sistema de garantia de qualidade como política de atendimento, bem como em ensaios clínicos. Concluíram que as intervenções de *feedback* para melhorar o resultados do tratamento, levaram a maior eficácia e efetividade nos atendimentos.

Segundo Nelson (2010) o paradigma de pesquisa mudou o foco da pesquisa centrada no paciente para os resultados do tratamento, passando dos resultados comparados com grupos e focalizou os resultados no nível do cliente individual. Este movimento envolve o monitoramento de resultados, para fins de planejamento de tratamento e cuidados de qualidade. Alguns desses sistemas de monitoramento incluem sistemas de alerta precoce, que poderia ajudar a identificar e atender melhor os clientes com risco de piora no tratamento ou desistência. O autor avaliou o atendimento de jovens, identificou as variáveis preditoras do andamento do tratamento, e a necessidade de criação de notas de corte utilizando o *Youth Outcome Questionnaire*-30 (YOQ-30), versão do *Outcome Questionnaire para adolescentes*. O estudo também replicou a precisão de um sistema de alerta para os clientes jovens em situação de risco, explorando várias abordagens para a criação de pontos de corte do sistema de alerta utiliza para suas previsões, e relatar a respectiva precisão de cada um. O estudo traz contribuições para futuros estudos comparando os resultados entre os grupos de clientes cujos terapeutas fazem ou não receber feedback sistemático. Este empreendimento oferece muitos benefícios para os esforços de melhoria de qualidade que estão sendo feitas por médicos e planos de saúde.

O monitoramento da resposta do paciente ao tratamento deveria ser integrado não só pelos terapeutas, mas também pelos supervisores de modo que o aluno já incluísse em suas primeiras experiências a avaliação contínua dos resultados, permitindo maximizar a recuperação dos pacientes que não respondem ao tratamento de modo esperado. A avaliação contínua dos resultados ajuda a concretizar os objetivos da supervisão (Worthen & Lambert, 2007).

Os resultados das pesquisas mostram que a psicoterapia, nos modelos de intervenção individual ou grupal, é um tratamento eficaz para uma variedade de condições seja por dificuldades psicológicas ou deficiências cognitivas, intelectuais, fisicas, auditivas e visuais. Segundo a Associação Americana de Psicologia (APA, 2013) a psicoterapia resulta em

benefícios para as pessoas que precisam dos serviços de saúde mental, assim deve ser incluída no sistema de cuidados da saúde.

OBJETIVOS

Objetivo Geral

O objetivo principal da presente pesquisa foi obter evidências de precisão e validade da versão brasileira do *Outcome Questionnaire* OQ 45.2.

Objetivos Específicos

Avaliar a precisão do OQ 45.2 pelo método de teste-reteste e pela consistência interna da escala total e subescalas com o alfa de Cronbach, em relação aos itens que as compõem.

Verificar a validade de critério por grupos contrastantes, comparando pacientes e não pacientes.

Verificar se existem diferenças entre os resultados de homens e mulheres em relação às médias do OQ 45.2.

Verificar a validade relativa à consistência interna entre o resultado global e as subescalas.

Obter evidências de validade concorrente com os seguintes instrumentos: Escala de Avaliação de Sintomas (EAS-40), Inventário de Depressão de Beck (BDI II) e Questionário Geral de Saúde (QSG).

MÉTODO

Participantes

A amostra foi constituída por 419 participantes adultos e obtida por conveniência, subdividida em dois grupos. Um grupo constituído por 59 pacientes adultos, com 18 anos ou mais, atendidos em uma Clínica-Escola de Psicologia, situada em uma cidade da Grande São Paulo. O outro grupo teve 360 participantes adultos, com 18 anos ou mais, que será doravante denominado de amostra não clínica ou grupo não-pacientes. Os participantes do grupo de pacientes, que procuraram a Clínica-escola no período de outubro de 2011 a outubro de 2012 e aceitaram participar da pesquisa, refletem a população real e não uma amostra estratificada. Foram convidados nas entrevistas de triagens e as aplicações das escalas foram realizadas pela pesquisadora.

Os participantes da amostra não clínica foram contatados em locais destinados ao estudo, trabalho e culto religioso, sendo esse grupo formado pelos que aceitaram participar da pesquisa. As características demográficas, de acordo com gênero, escolaridade, idade e estado civil, da amostra total são apresentadas na Tabela 1.

Tabela 1. Dados demográficos da amostra total (N = 419)

Variável	Descrição	Frequência	%
Sexo	Masculino	156	37,2
	Feminino	263	62,8
Escolaridade	Fundamental	34	8,0
	Médio	162	39,0
	Superior	193	46,0
	Não responderam	28	7,0
Idade	18-19	52	12,4
	20-39	181	43,2
	40-59	154	36,7
	60-78	28	6,7
	Não responderam	4	1,0
Estado Civil	Solteiro	246	58,7
	Casado/União estável	117	27,9
	Divorc./Separ./Viúvo	34	8,1
	Não responderam	22	5,3

A distribuição entre os gêneros não foi equilibrada na amostra total, sendo 37,2% foi de homens e 62,8 % de mulheres, porque nos locais onde foram coletados os dados havia maior número de mulheres que concordaram em colaborar. Na amostra total 46% dos participantes tinham nível de escolaridade superior incompleto ou completo e 39% nível médio incompleto ou completo. Quanto à faixa etária, 43,2% tinham idade entre 20 e 29 anos, 36,7% entre 30 e 59 anos, 12, 4% entre 18 e 19 anos e apenas 6,7% com 60 anos ou mais. Em relação ao estado civil, 397 participantes informaram a sua condição e encontram-se assim distribuídos: 58,7% são solteiros, 27,9% casados ou em união estável, 8,1% viúvos, separados ou divorciados e 5,3% não responderam.

Na Tabela 2 os dados sociodemográficos são apresentados de acordo com os grupos de pacientes e não pacientes.

Tabela 2. Dados demográficos dos grupos de paciente e não pacientes

Variável	**Descrição**	**Pacientes**		**Não Pacientes**	
		N	**%**	**N**	**%**
Sexo	Masculino	22	37,2	134	37,2
	Feminino	37	62,7	226	62,7
Escolaridade	Fundamental	11	18,6	22	6,1
	Médio	24	40,7	141	39,2
	Superior	19	32,2	174	48,4
	Não responderam	5	8,5	23	6,4
Idade	18-19	2	3,4	50	14,0
	20-39	29	49,1	218	60,5
	40-59	20	34,0	66	18,3
	60-78	8	13,5	22	6,1
	Não responderam	0	0	4	1,1
Estado Civil	Solteiro	21	35,6	225	62,5
	Casado/União est.	28	47,4	117	32,5
	Divorc./Separ./Viúvo	10	17,0	34	9,4
	Não responderam	0	0	22	6,1

Na amostra não clínica também a proporção de mulheres que concordaram em participar da pesquisa foi maior do que a dos homens, tendo sido mantida a mesma proporção de participantes da amostra clínica e da global. O desequilíbrio na amostra clínica ocorreu

porque a demanda de mulheres na clínica-escola foi maior. Em outras pesquisas de caracterização sociodemográfica da clientela de clínica-escola foram encontradas frequências semelhantes (Enéas, Faleiros & Sá, 2000; Campezatto & Nunes, 2007).

Quanto à escolaridade observa-se que o grupo de pacientes apresenta uma proporção maior de participantes com Ensino Fundamental (18,6%) do que o de controle (6,2%) e no Ensino superior a proporção se inverte, com uma porcentagem maior do grupo controle (48,4%) do que no de pacientes (32,2%). As porcentagens relativas ao Ensino Médio são semelhantes nos dois grupos (40,7% dos pacientes e 39,2% o dos não-pacientes).

A distribuição da amostra de pacientes por faixa etária é similar às faixas do estudo original (Lambert et al., 2004) para que os resultados das duas pesquisas possam ser comparados. A amostra de pacientes da amostra brasileira teve as seguintes proporções: < 20 anos (3,4%), 20-39 anos (49,1%), 40-59 anos (34,0%) e 60 ou mais (13,5%). No estudo americano para verificar se havia diferenças nas médias entre as faixas etárias, eles trabalharam com amostra de 504 pacientes, sendo faixa < 20 anos (4,2%); de 20-39 anos (60,1%); de 40-59 anos (34,1%) e 60 anos ou mais (1,5%).

Nas duas amostras de pacientes, brasileiros (49,1%) e americanos (60,1%), houve maior concentração na faixa entre 20 e 39 anos, embora maior na americana. Outra diferença apareceu nos maiores de 60 anos, sendo 13,5% na amostra brasileira e somente 1,6% na americana. Essa diferença pode ser devida a sua origem, no caso da americana os pacientes eram do EAP (*Employee Assistance Program*), que é um programa de assistência ao funcionário gerenciado pela *Human Affairs International*, enquanto que na brasileira foram os que procuraram uma clínica-escola para atendimento psicológico.

Em relação à amostra não clínica a maior concentração de participantes com idades entre 20 e 39 anos está relacionada aos locais de coleta de dados uma universidade particular e um curso técnico localizados na grande São Paulo.

Quanto ao estado civil, a maioria dos não pacientes era composta por solteiros (62,5%), enquanto nos não pacientes a maioria era casada (40,7%). A frequência maior de solteiros provavelmente está relacionada aos locais em que os dados foram coletados, curso técnico e universidade, o que não acontece no grupo clínico.

Quanto à ocupação atual declarada pelos participantes, a amostra total (N = 419) foi distribuída da seguinte forma: 61,3% estão trabalhando, 16,5% estão desempregados, 2,6% em beneficio, 4,8% são aposentados, 5,3% nunca trabalharam e, 9,5% não responderam. Os dados são apresentados na Tabela 3.

Tabela 3. Distribuição dos grupos de pacientes e não pacientes de acordo com a ocupação atual

Ocupação atual	Amostra					
	Não pacientes		Pacientes		Total	
	N	%	N	%	N	%
Trabalhando	234	65,0	23	39,0	257	61,3
Desempregado	53	14,7	16	27,1	69	16,5
Em benefício	6	1,7	5	8,5	11	2,6
Aposentado	16	4,4	4	6,8	20	4,8
Estudando	19	5,3	3	5,1	22	5,2
Não responderam	32	8,9	8	13,6	40	9,5
Total	360	86,0	59	14,0	419	100

No grupo de pacientes 39% estão atualmente trabalhando, 27,1 estão desempregados, 8,5% recebendo benefício por doença. Enquanto no grupo de não pacientes 65% dos participantes estão trabalhando, 14,7 desempregados, somente 1,7% recebem benefício por doença.

A amostra também foi descrita quanto aos recursos econômicos. Para a obtenção de dados quanto à classe econômica dos participantes foi utilizado o Critério de Classificação Econômica Brasil da Associação Brasileira de Empresas de Pesquisa (ABEP), que é um critério que tem como função estimar o poder de compra das pessoas. Utiliza um sistema de pontos, identificando e pontuando a posse de itens tais como TV, rádio, automóvel, número de banheiros na residência, máquina de lavar, geladeira, freezer, entre outros e o grau de escolaridade do chefe da família. A classificação de classes econômicas da ABEP é dividida em oito classes e cada uma corresponde uma renda média familiar (valor bruto em R$): A1 = R$ 12,926,00; A2= R$ 8.418,00; B1= R$ 4.418,00; B2= R$ 2.565,00; C1= R$ 1.541,00; C2= R$ 1.024,00; D=R$ 714,00; e E=R$ 477,00. Na Tabela 4 é apresentada a classificação da amostra.

Tabela 4. Classificação da classe econômica da amostra segundo o Critério de Classificação Econômica Brasil da ABEP

Classe	Renda Bruta R$	Não pacientes		Pacientes		Total	
		N	%	N	%	N	%
A1	12.926,00	0	0	0	0	0	0
A2	8.418,00	19	5,3	0	0	19	4,5
B1	4.418,00	40	11,1	8	13,6	48	11,4
B2	2.565,00	128	35,6	20	33,9	148	35,2
C1	1.541,00	97	26,9	10	16,9	107	25,5
C2	1.024,00	43	11,9	10	16,9	53	12,6
D	714,00	8	2,2	5	8,5	13	3,1
E	477,00	1	0,3	1	1,7	2	0,5
Não resp.		24	6,7	5	8,5	29	7
Total		360	85,9	59	14,1	419	100

Na classificada da ABEP cada faixa corresponde à Renda familiar média – Valor Bruto em Reais, com base no Levantamento Sócio Econômico de 2010. Na amostra total, 35,2% classificam-se na Classe B2; 25,5% na C1; 12,6% na C2; 11,4% na B1; 4,5% na A2; 3,1% na D; e 0,5% na E. Ao comparar os grupos de pacientes e não pacientes tendo como referência a renda familiar, a classe econômica A2 só aparece na amostra não clínica, 33,9% dos pacientes e 35,6% do grupo de não pacientes se classificam na B2. Quanto à classe econômica C1, observa-se 26,9% grupo de não pacientes e 16,9% da amostra clínica. De modo geral, as amostras apresentam maior proximidade em relações ao poder de compra na classe B2. As maiores diferenças apareceram nas classes A2 e D, no grupo de pacientes estudado nenhum foi incluído na classe A2, enquanto na classe D observa-se 8,5% dos pacientes.

Instrumentos

Foram utilizados os seguintes instrumentos:

Escala de Avaliação de Resultados OQ - 45.2

Escala de Avaliação de Sintomas – ESA-40

Questionário de Saúde Geral de Goldberg – QSG

Inventário de Depressão de Beck – BDI II

Termo de Consentimento – grupo pacientes (Anexo A)

Termo de Consentimento – grupo não pacientes (Anexo B)

Escala de Avaliação de Sintomas – EAS-40 (Laloni, 2001)

A EAS-40 é uma escala do tipo autorrelato de 40 itens respondidos em uma escala *Likert* de três pontos, em que zero corresponde a nenhum sintoma, um ponto indica um pouco do sintoma e dois pontos, muitos sintomas. Os 40 itens são distribuídos em quatro dimensões, com 10 itens cada: Psicotismo, Obsessividade-compulsividade, Somatização e Ansiedade. A escala avalia assim a intensidade de sintomas psicopatólogicos e foi adaptada do Symptom Checklist-90-R (SCL-90-R) por Laloni (2008).

Inventário de Depressão de Beck – BDI-II (Gorenstein, et al. 2011)

O Inventário de Depressão de Beck (BDI-II) é um instrumento de autorrelato, com o objetivo de medir a intensidade da depressão em adultos e adolescentes. A versão II avalia os sintomas correspondentes aos critérios diagnósticos dos transtornos depressivos descritos no DSM-IV, e a adaptação brasileira foi realizada por Gorenstein, Pang, Argimon e Werlang (2011). É um questionário com 21 grupos de afirmações, em que cada item é avaliado segundo uma escala de quatro pontos que varia de 0 a 3. Os sintomas estão agrupados em duas dimensões: Somato-afetiva e Sintomas de natureza psicológica e cognitiva.

Questionário Geral de Saúde de Goldberg – QSG (Pasquali, et al., 2011)

Constitui um questionário com 60 itens sobre sintomas psiquiátricos não psicóticos, a ser respondido em uma escala do tipo *Likert* de quatro pontos. Uma parte dos itens expressa sintomas de forma direta, e a outra se refere a comportamentos normais. Quando o item é formulado como sintoma, a escala tem o seguinte sentido: 1= não, absolutamente, 2= não mais do que de costume, 3= um pouco mais do que de constume e 4= muito mais do que de costume. Quando indica comportamento normal, a escala é invertida, do seguinte modo: 1= melhor que de costume, 2= como de costume, 3= pior do que de

costume e 4= muito pior do que de costume (Pasquali et al.1996). Os itens foram agrupados em 5 dimensões: Stress Psíquico; Desejo de morte; Desconfiança no próprio desempenho, Distúrbio do sono; Distúrbios psicossomáticos, e Saúde Geral (soma total dos escores dividido por 60, e mede a severidade da ausência de saúde mental).

Procedimento

O Projeto desta Pesquisa foi aprovado pelo Comite de Ética da Universidade Guarulhos (CAAE 0041.0.132.000-110). Os instrumentos foram aplicados nos participantes, sendo que para todos foi entregue o Termo de Consentimento Livre e Esclarecido no qual constavam os objetivos da pesquisa e a forma de divulgação dos resultados de acordo com as normas éticas. Somente participaram da pesquisa os participantes que concordaram com os procedimentos da pesquisa e assinaram o TCLE. Além disso, a coleta dos dados implicou sempre a presença da pesquisadora, possibilitando que possíveis dúvidas dos participantes fossem esclarecidas.

Todos os participantes responderam ao menos um dos instrumentos da pesquisa, os dados são apresentados na Tabela 5.

Tabela 5. Total de testes respondidos pela amostra total e por sexo.

Instrumentos	Mulheres	Homens	Participantes
EAS-40	239	134	373
BDI-II	123	66	189
QSG	75	40	115
Nº de instrumentos respondidos pelos participantes (N=419)			
OQ 45.2			23
OQ 45.2/ EAS-40			115
OQ 45.2/ BDI-II			01
OQ 45.2/QSG			21
OQ 45.2/EAS-40/BDI-II			165
OQ 45.2/EAS-40/QSG			71
OQ 45.2/BDI-II/QSG			01
OQ 45.2/EAS-40/BDI-II/QSG			22
Total			**419**

Na amostra total 23 participantes responderam somente o OQ 45.2, deixando em branco os outros protocolos, 137 participantes responderam dois dos quatro instrumentos, 237 responderam três instrumentos e 22 participantes responderam todos os instrumentos

utilizados na pesquisa. Em relação à coleta de dados, no caso de universitários, estudantes de curso técnico, programa da juventude e idosos da universidade aberta, as aplicações foram realizadas em salas de aulas nos seus locais de estudo. Em horário previamente combinado com os professores, a pesquisadora se apresentou às classes, informou sobre os objetivos da pesquisa, esclareceu tratar-se de participação voluntária e a garantia de sigilo da identidade dos participantes. Além disso, foi informado sobre a possibilidade de uma segunda aplicação de um dos instrumentos com fins de reteste, no intervalo de 7 a 14 dias. No caso dos pacientes a aplicação foi individual, em salas de atendimento de uma clínica-escola da grande São Paulo, pela pesquisadora.

Um subgrupo da amostra foi obtida nos locais de estudo, trabalho e culto religioso dos participantes, algumas pessoas foram convidadas pelos próprios participantes da pesquisa e concordaram em colaborar com a pesquisa, nestes casos as aplicações ocorreram no consultório da pesquisadora.

A amostra abrangeu os seguintes subgrupos: a) alunos de graduação do 1º e 2º semestres de uma universidade particular da grande São Paulo, sendo: 6,4% de Educação Física, 8,8% de Enfermagem e, 6,7% de Nutrição; b) alunos de cursos de graduação Tecnológicos, a saber: 3,8% de Gestão de Qualidade, 14,6% de Radiologia e, 6% de Recursos Humanos; c) 8,6% de alunos oriundos de um programa preparatório para o ingresso no mercado profissional (Programa da Juventude); d) 5,3% de idosos com mais de 60 anos que frequentavam cursos livres para a terceira idade em uma universidade particular da grande São Paulo; e) 7,4% de pessoas indicadas por outros participantes, f) 1,7% voluntários de uma Igreja Católica da região leste da cidade de São Paulo e g) 14,1% de pacientes de uma clínica-escola da grande São Paulo.

A amostra para obtenção do reteste foi composta por 55 participantes, sendo 29 do sexo masculino e 26 do sexo feminino, com um intervalo de sete a 14 dias. O reteste dos pacientes ocorreu durante a segunda entrevista para fechamento ou encaminhamento do paciente.

Para obtenção da coerencia interna entre os itens (alfa de Cronbach) e as correlações entre as subescalas e o resultado total do OQ 45.2 foi utilizada a amostra total (N= 419).

Para a validade concorrente os resultados obtidos na avaliação com a Escala de Avaliação de Resultados de Psicoterapia OQ-45.2 foram correlacionados com os da Escala de Avaliação de Sintomas (EAS-40), do Inventário de Depressão de Beck (BDI II) e do Questionário de Saúde Geral de Goldberg (QSG). Os três instrumentos foram utilizados como critérios externos, sendo que para cada um deles foi empregado um subgrupo da

amostra total, descritas a seguir: A Escala de Avaliação de Sintomas- EAS-40 foi aplicada a 373 participantes, sendo 134 homens e 239 mulheres, destes 51 do grupo de pacientes e 95 do grupo de não pacientes; o Inventário de Depressão de Beck (BDI-II), a 189 participantes, sendo 66 homens e 123 mulheres, 22 pacientes e 167 não pacientes; e o Questionário Geral de Saúde (QSG) a 115 participantes, sendo 40 homens e 75 mulheres, 20 do grupo de pacientes e 95 de não pacientes.

Para determinar se todos os itens estão avaliando no mesmo sentido que o resultado total, foi calculado a correlação item-total e a média de pontos de cada item.

RESULTADOS E DISCUSSÃO

Precisão de Teste-Reteste

Os critérios para interpretação das correlações foram extraídos do texto de Bisquera, Sarriera e Martínez (2004), que estabelece os seguintes valores para a interpretação da Correlação de *Pearson*: r = 1,00 perfeita; 0,80 ≤ r <1 muito alta; 0,60 ≤ r < 0,80 alta; 0,40 ≤ r < 0,60 moderada; 0,20 ≤ r < 0,40 baixa; 0< r < 0,20 muito baixa; e r = 0 nula, não existe relação entre os dados, as variáveis são independentes. Para os coeficientes negativos, a interpretação é idêntica.

A precisão de Teste-Reteste foi feita com 55 participantes, os coeficientes de correlação de Pearson variaram entre 0,89 e 0,75, como se segue: 0,89 (OQ total), 0,88 (Desconforto Subjetivo - SD), 0,79 (Relações Interpessoais - IR) e 0,75 (Papel Social - SR), indicando boa estabilidade temporal tanto da escala total quanto em relação às suas três dimensões. Todas as correlações entre teste e reteste foram significantes (p≤0,01) para o escore total do OQ-45.2, e as suas três dimensões/subescalas, sendo que as duas primeiras podem ser consideradas muito altas e as demais altas. Esses resultados mostram que o OQ-45.2 apresenta boa estabilidade temporal, tanto para o escore total como para as três subescalas, uma vez que foram obtidos coeficientes acima de 0,70.

O reteste foi realizado em um intervalo entre as aplicações que variam de sete a 14 dias, para apenas parte da amostra (n = 55), destes foram retirados os dados dos estudantes universitários (n = 33) para fins de comparação com o estudo de Lambert, et al. (2004) e Jong et al. (2007). Para comparação dos resultados desta pesquisa com os da literatura são apresentados na Tabela 6 os estudos de Precisão Teste-Reteste realizados com amostras de estudantes universitários dos Estados Unidos (E.U.A), Holanda e Brasil (presente estudo). Nos estudos de Lambert et al. (2004) e Jong et al. (2007) a precisão Teste-Reteste foi realizado somente com amostra de estudantes universitários do grupo controle. Na presente pesquisa foi realizado um estudo de precisão teste-reteste com um subgrupo de estudantes universitários (n = 33) que fazem parte da amostra de não-pacientes. Também há uma pequena diferença no intervalo de dias entre as duas aplicações nos três estudos.

Tabela 6. Comparação da Precisão pelo Reteste do OQ 45.2 em amostras de estudantes universitários (grupo não clínico)

	Americana*	Holandesa**	Brasileira***
SD	0,78	0,81	0,91
IR	0,80	0,71	0,83
SR	0,82	0,73	0,58
Total	0,84	0,82	0,89
N	157	264	33
Intervalo (dias)	21	14	7

*Lambert, et al. (2004) ** Jong et al. (2007) *** Silva (2013)

Foram encontradas correlações muito altas para OQ Total nos três estudos (entre 0,82 a 0,89), para a subescala Desconforto Subjetivo (SD) nas amostras holandesa e brasileira (respectivamente 0,81 e 0,91), para as subescalas (IR) nas amostras americana e brasileira (0,80 e 0,83) e para a subescala papel Social (SR) na americana (0,82). As correlações foram altas para SD na americana (0,78), e para IR e SR na holandesa (0,71 e 0,73) e moderadas para SR na brasileira (0,58). No subgrupo de estudantes universitários as correlações foram muito altas, exceto para a subescala SR que mostrou uma moderada estabilidade temporal (r = 0,58). Apesar de o reteste ter ocorrido em um intervalo de tempo menor ao ser comparado com os outros dois estudos, deve ser lembrado que o número de estudantes universitários, que responderam a escala em dois momentos diferentes, no presente estudo é bem menor, enquanto nas outras amostras o tamanho da amostra é menor que as outras e mais heterogênea, o que pode diminuir a correlação.

Considerando as pesquisas de Precisão Teste-reteste do OQ 45.2, em diferentes países, têm apontado correlações significativas. No Chile, La Parra e Bergen (2002) obtiveram uma correlação teste-reteste para OQ Total foi de 0,82, semelhante às obtidas nos estudos americano e holandês, mas menor que a da presente pesquisa, o resultado pode estar relacionado ao tamanho da amostra de universitários utilizados para o reteste. Os resultados da fidedignidade pelo método de teste-reteste demonstraram que a versão brasileira do OQ-45.2 apresentou precisão satisfatória na amostra estudada, exceto para a subescala SR, uma vez que esta é menor do que 0,60 valor mínimo aceito pelos critérios da Resolução do CFP para a precisão.

Consistência Interna

A consistência interna é um coeficiente de correlação entre os distintos itens do instrumento que determina se na base dos itens está o mesmo conceito, ou seja, que todos os itens medem o mesmo constructo. A consistência interna foi estudada pela homogeneidade dos itens representativos dos constructos de cada dimensão, pelo alfa de Cronbach, que é proposto como o mais adequado para instrumentos com escores múltiplos, em que as respostas aos itens variam numa escala graduada de pontos, neste caso de 0 a 4.

Na amostra geral os índices variaram entre 0,63 a 0,92 relacionados ao escore total e as dimensões Desconforto Subjetivo (SD), Relações Interpessoais (IR) e Papel Social (SR). Os valores obtidos indicam que os itens avaliam medem constructos semelhantes.

Na Tabela 7 podem ser observados os alfas de Cronbach obtidos para amostra total, estudantes e pacientes brasileiros e os da pesquisa de Lambert et al (2004), com estudantes universitários e pacientes americanos, para fins de comparação.

Tabela 7. Coeficientes de Alfa para OQ 45 Total e cada subescalas nas amostras brasileira e americana de estudantes universitários e pacientes

		Brasileira		Americana		
	Total	Universitários	Pacientes	Universitários	Pacientes	
	Itens				Alfa	N
SD	25	0,91	0,90	0,92	0,91	298
IR	11	0,71	0,64	0,74	0,74	294
SR	9	0,65	0,68	0,70	0,71	295
Total	45	0,93	0,92	0,93	0,93	289
N		156	59	157		

O coeficiente alfa do OQ-45.2 total na amostra brasileira considerando apenas os universitários foi de 0,93 e para os pacientes de 0,92. Estes coeficientes são semelhantes aos obtidos na amostra americana, para os universitários e pacientes (0,93). Somente para a subescala SR os coeficientes foram mais baixos na amostra brasileira (0,65 e 0,68) do que os da americana (0,70 e 0,71), nos dois grupos, embora essa subescala também tenha apresentado as menores correlações nas amostras americanas. O alfa de Cronbach para a amostra chilena de La Parra e Bergen (2002) para a pontuação total do OQ-45.2 foi de 0,91,

semelhante ao das amostras brasileira e americana.

O coeficiente alfa obtido para a subescala IR, no grupo de pacientes, é menor (0,64) do que o obtido nos dois grupos americanos (0,74) e nos universitários brasileiros (0,71). Para a subescala SR o alfa no grupo de universitários brasileiros (0,65) é menor do que o dos americanos (0,70). Nos grupos de universitários e pacientes, o coeficiente de alfa obtido para o escore total e a subescala SD são similares aos da amostra americana, indicando que o OQ-45.2 tem boa consistência interna para esses dois resultados.

Os coeficientes encontrados por Le Coco et al. (2008), referentes à amostra italiana, foram 0,93 para escore global, 0,92 para SD, 0,77 para IR e 0,66 para SR, sendo muito similares aos da amostra de não pacientes americana, exceto para a escala SR, que é mais próxima dos resultados das amostras brasileiras. Na Tabela 8 são indicados os dados comparativos para as amostras, brasileira e americana, de pacientes e não pacientes.

Tabela 8. Coeficientes Alfa de Cronbach para OQ 45.2 e para cada subescala das amostras brasileira e americana de pacientes e não pacientes

		Brasileira		Americana	
	Total Itens	Paciente*	Não Paciente**	Paciente	Não Paciente
SD	25	0,90	0,88	0,91	0,92
IR	11	0,64	0,67	0,74	0,74
SR	9	0,68	0,59	0,71	0,70
Total	45	0,92	0,90	0,93	0,93

*59 pacientes ** 360 participantes

O coeficiente alfa do escore total do OQ-45.2 para o grupo de pacientes brasileiros foi 0,92 e 0,93 para os pacientes americanos. Para os grupos de não pacientes brasileiros o coeficiente alfa foi 0,90 e 0,93 para o grupo de não pacientes americanos. Entretanto para as subescalas IR nos grupos de pacientes e não pacientes o alfa foi menor em relação aos grupos americanos. Na subescala SR observa-se a maior diferença, o coeficiente alfa foi menor no grupo não clínico da amostra brasileira 0,59 enquanto que no grupo de não pacientes americano foi de 0,70.

No Chile, Rodriguez (2000) investigou a consistência interna para o OQ-45.2 e as três subescalas de uma amostra composta por dois grupos, um de pacientes (N = 124) e não paciente (N = 202). Foram obtidos os seguintes coeficientes alfa: para o escore total 0,91 em

ambos os grupos; para SD 0,89 e 0,88 (respectivamente); para IR 0,67 e 0,71 e para SR 0,63 e 0,65. Esses resultados foram muito semelhantes aos da amostra brasileira, também com coeficientes menores para escala SR.

Assim em todas as pesquisas citadas anteriormente pode ser constatado que o coeficiente de consistência interna da subescala SR é mais baixo que as demais, provavelmente porque essa escala tem um número de itens muito menor do que o das outras. Dessa forma coeficiente obtido na amostra brasileira de não pacientes (0,59) está de acordo com a tendência observada nas pesquisas internacionais.

Validade

Correlações entre as subescalas

Para avaliar as relações entre as pontuações das subescalas e a escala total do OQ 45.2 foram calculadas e são apresentadas na Tabela 9.

Tabela 9. Correlação entre pontuação total e as subescalas na amostra total (N=419)

Dimensões	Total	SD	IR
SD	0, 95*		
IR	0, 78*	0,70*	
SR	0, 67*	0, 58*	0, 46*

*$p \leq 0,01$

Os resultados indicaram correlação muito alta entre o escore total e a subescala SD (0,95), correlações altas com IR (0,78) e SR (0,67), assim como entre SD e IR (0,70). Entre as subescalas SD e SR (0,58) e IR e SR (0,46) foram encontradas correlações moderadas. Esses dados mostram que a subescala SR é que tem menor relação com o resultado global e com as demais subescalas, provavelmente as diferenças estejam relacionadas ao número menor de itens da subescala SR (nove itens).

Tabela 10. Correlações entre as subescalas SD, IR, SR e escore total para cada grupo

	Total		SD		IR	
Grupos	Paciente	Não Paciente	Paciente	Não Paciente	Paciente	Não Paciente
SD	0,91*	0,96*				
IR	0,73*	0,76*	0,70*	0, 65*		
SR	0,69*	0, 66*	0,71*	0,51*	0,56*	0,40*

* $p \leq 0,01$

As correlações foram muito altas entre escore total do OQ 45.2 e SD nos grupos de pacientes e não-pacientes (respectivamente 0,91 e 0,96), foram altas entre o escore total, IR (0,73 e 0,76) e menores com SR (0,69 e 0,66), entre SD e IR (0,70 e 0,65), entre SD e SR (0,71 e 0,51). As correlações entre IR e SR (0,56 e 0,40) foram baixas nos dois grupos. As correlações mais baixas entre as subescalas podem ter ocorrido em função do menor número de itens das mesmas.

No estudo realizado no Chile (Rodriguez, 2000) foram obtidos os seguintes valores de consistência interna para o OQ 45.2 e as três subescalas. Foram obtidos os seguintes coeficientes alfa nos grupos de pacientes (N = 124) e não paciente (N = 202) para escore total 0,91 em ambos os grupos; para SD 0,89 e 0,88 (respectivamente); para IR 0,67 e 0,71 e para SR 0,63 e 0,65. Esses resultados foram muito semelhantes aos da amostra brasileira, também com coeficientes menores para escala SR.

No estudo americano (Lambert et al. 2004) apresentou uma maior consistência interna nas duas amostras, no grupo de pacientes os valores obtidos para escore total, SD, IR e SR, respectivamente 0,93, 0,91, 0,74 e 0,71 e para a comunidade foram 0,93, 0,92, 0,74 e 0,70, o que indica que os itens da versão americana mostraram maior consistência interna.

Grupos Contrastantes

Para a validade de grupos contrastantes, os resultados do OQ-45.2 foram estudados em um grupo de pacientes (N = 59) e outro de não pacientes (N = 360).

Nota-se que há uma diferença entre as médias e os desvios padrões nos grupos de pacientes e não-pacientes. Para determinar se as diferenças entre as médias são significativas foram calculadas a nota *t*, os dados são apresentados na Tabela 11.

Tabela 11. Médias, desvios padrão e teste *t* entre os grupos de pacientes e não pacientes

Grupos	Não Pacientes (N = 360)		Pacientes (N = 59)		*t*	*Sig*
	Média	DP	Média	DP		
OQ Total	49,81	20,55	79,86	32,71	6,84**	0,000
SD	26,29	13,16	45,64	18,48	7,73**	0,000
IR	12,68	6,26	18,71	7,34	5,97**	0,000
SR	10,93	4,44	13,44	6,53	2,85*	0,006

** p< 0,001; *p< 0,01

Conforme o esperado, os resultados apontaram médias inferiores para o grupo de não pacientes. De acordo com os valores de *t*, as diferenças entre os dois grupos são significantes e indicam a sensibilidade do OQ-45.2 para diferenciar grupos de pacientes de não pacientes, mostrando que o instrumento é capaz de discriminar os grupos. O instrumento tem uma boa capacidade para discriminar entre distintos grupos clínicos de não-clínicos.

No estudo realizado por Lambert et al. (2004) a média do escore total do OQ-45.2 do grupo de pacientes da EAP (N= 441) foi 73,61, o DP 21,39 e o *t* = 6,05, enquanto no presente estudo com amostra brasileira de pacientes a média foi 79,86, o DP 32,71 e *t =6,84*. Na amostra brasileira de não pacientes (N= 360) a média para escore total foi 49,81 e o DP = 20,55, na amostra americana (Lambert et al., 2004) a média dos não pacientes (universitários) foi 46,49, o DP 19,82 e o *t* = 1,15, enquanto na amostra da comunidade (não pacientes) a média foi 45,19, o DP 18,57 e o *t* = 24,52.

Os dados obtidos na amostra brasileira indicam que a adaptação do OQ-45.2 é capaz de discriminar grupos clínicos de não-clínicos.

Considerando que o tamanho das amostras costuma exercer grande influência sobre os coeficientes de correlação, os resultados das amostras feminina e masculina foram calculados separadamente. Na Tabela 12 são apresentadas as médias e desvio padrão da amostra total em relação aos gêneros, bem como os resultados dos testes *t* relativos a esta variável.

Tabela 14. Diferenças entre gêneros nas Médias e Desvio Padrão para o OQ Total e as subescalas

OQ-45.2	Homens (N= 156)		Mulheres (N= 263)		Teste *t*	
	Média	DP	Média	DP	*t*	*Sig.*
SD	24,13	13,40	31,90	16,02	-5,327**	0,000
RI	11,97	5,83	14,44	7,08	-3,863**	0,000
SR	10,57	4,76	11,71	4,87	-2,341*	0,020
Total	47,33	23,18	58,03	25,17	-4,430**	0,000

** p < 0,001; *p < 0,05

No presente estudo, foram encontradas diferenças significantes quanto ao gênero, sendo que as mulheres apresentaram médias maiores do que os homens para o resultado total e para cada subescala. Na pesquisa holandesa também foram observadas diferenças por gênero tanto na amostra clínica e como na de não pacientes. Jong et al. (2007) observou na amostra holandesa, que os homens obtiveram maior pontuação na subescala SR que refere-se ao desempenho do papel social, enquanto que as mulheres mostraram maiores pontuações na subescala SD.

Os resultados desta pesquisa apontaram diferenças significantes entre os dois grupos para o escore total (p <0,001) e para as subescalas, sendo as médias das mulheres são maiores do que a dos homens. As mulheres relatam maior percepção do sofrimento psíquico, presença de sentimentos e sensações interiores que causam desconforto subjetivo, o que justificar a maior procura de terapia por parte das mulheres.

Na pesquisa Chilena, Rodríguez (2000) também encontrou diferenças significativas ao comparar o escore total e das subescalas do OQ-45.2 em amostra da comunidade, as mulheres tem médias maiores que os homens para SD, IR, SR e escore total, a autora relaciona tais diferenças a fatores culturais, 75% das mulheres buscam psicoterapia, e que os homens por trabalhar fora de casa e, portanto teriam mais dificuldades de horário e menos acesso a psicoterapia. Entretanto essa diferença não ocorreu na pesquisa de Lambert et al. (2004), pois ele relata que não foram observadas diferenças significativas quanto ao gênero.

Validade Concorrente

Escala de Avaliação de Sintomas (EAS-40)

Inicialmente será relatada a comparação com a Escala de Avaliação de Sintomas (EAS-40). A Tabela 13 apresenta os coeficientes de correlação de Pearson dos escores do OQ-45.2 com a Escala de Avaliação de Sintomas (EAS-40). A EAS-40 tem 4 dimensões: Psicotismo, Obsessividade-Compulsividade, Somatização e Ansiedade.

Tabela 13. Correlação entre os escores total e as dimensões da EAS-40 com escore total e dimensões do OQ-45.2 (NP = 322, P = 51).

OQ 45.2	Escala de Avaliação de Sintomas (EAS-40)				
	Total	Somatiz.	Obs.Com.	Psicotismo	Ansiedade
Total	0,80*	0,56*	0,68*	0,75*	0,75*
SD	0,81*	0,57*	0,70*	0,75*	0,78*
IR	0,59*	0,42*	0,50*	0,55*	0,56*
SR	0,53*	0,34*	0,49*	0,49*	0,48*

* p < 0,001

NP = Não pacientes; N = Pacientes.

A correlação entre os totais das duas escalas foi de 0,80. Nota-se que as correlações dos escores total do EAS-40 com as subescalas SD, IR, SR do OQ-45.2 variaram de 0,53 a 0,81. Todas as correlações entre as duas escalas foram significantes (p < 0,001). As correlações foram altas entre o escore total do OQ-45.2 e das escalas Psicotismo e Ansiedade (0,75) e menores com Somatização (0,56) e Obsessividade e compulsividade (0,68). A subescala SD teve correlações semelhantes com as escalas da EAS-40, com resultado menor para Somatização (0,57).

O resultado total da EAS-40 mostrou maior correlação com a subescala Desconforto Subjetivo (SD = 0,81), o que indica uma relação maior entre esta dimensão e o total da EAS-40, que avalia a presença de sintomas psicológicos e psicopatológicos enquanto a subescala de Desconforto Subjetivo relaciona-se a transtornos de humor e ansiedade. Já as correlações do escore total da EAS-40 foram 0,59 com IR e 0,53 com (SR), indicando que elas provavelmente abrangem aspectos diferentes da personalidade. As subescalas IR e SR apresentaram correlações baixas (entre 0,34 e 0,56) com as dimensões da EAS-40, mostrando que elas avaliam poucos aspectos comuns entre elas.

No estudo de Lambert et al. (2002), com uma amostra de estudantes, foi utilizada a SCL-90-R (da qual a EAS-40 é derivada), para obtenção da validade concorrente do OQ-45, sendo obtidas as seguintes correlações: SD (0,61) e OQ Total (0,78), alcançando um nível de significância de p <0,01 com uma versão preliminar com 43 itens do OQ. Não constam no

manual do teste os índices de validade para IR e SR. No estudo de validade com amostra alemã (Lambert, Hannover et al, 2002), foram encontradas as seguintes correlações: para SD = 0,76, IR = 0,53, SR = 0,47 e escore total = 0,73.

No estudo brasileiro ao invés de utilizar o SCLR-90-R foi aplicada a Escala de Avaliação de Sintomas (EAS-40), resultado de um estudo de validade e precisão com a SCLS-90-R e no processo de adaptação a escala foi nomeada de EAS-40 realizado em uma amostra brasileira (Laloni, 2001). Os resultados referentes à validade concorrente entre o OQ 45.2 e EAS-40 indicam que estes instrumentos estariam medindo constructos semelhantes entre si. Outros estudos também foram encontrados indicativos de validade de constructo do OQ 45.2 com sintomas de depressão e ansiedade (Lambert,et al, 2004; Lo Coco et al., 2008; Jong, et al, 2007; Rodrigues, 2000).

Os resultados da validade do OQ-45.2 com o EAS-40 nos grupos de pacientes (N = 51) e não pacientes (N = 322) são apresentados na Tabela 14.

Tabela 14. Correlações entre a EAS-40 e o OQ-45.2 para os dois grupos Pacientes (N = 51) e não pacientes (N = 322)

	OQ-45.2 Total		SD		IR		SR	
EAS-40	P	NP	P.	NP	P	NP	P	NP
Escore total	0,78*	0,74*	0,83*	0,75*	0,64*	0,50*	0,65*	0,45*
Somatização	0,73*	0,47*	0,80*	0,48*	0,60*	0,33*	0,63*	0,25*
Obs-Comp	0,66*	0,61*	0,72*	0,62*	0,53*	0,41*	0,61*	0,35*
Psicotismo	0,76*	0,67*	0,78*	0,66*	0,62*	0,44*	0,53*	0,43*
Ansiedade	0,69*	0,72*	0,75*	0,75*	0,55*	0,50*	0,65*	0,38*

* $p < 0,001$
P = Pacientes; NP= Não Pacientes

Ao comparar dados separados das dimensões do EAS-40 com as subescalas do OQ 45.2, nos grupo de pacientes e não pacientes, observa-se que as correlações mais altas (maiores que 0,70) foram do escore total da EAS-40 e suas quatro dimensões no grupo de pacientes. Enquanto no grupo não-clínico, as correlações escore total do EAS-40 e três da suas quatro dimensões (maior do que 0,60, exceto Somatização menor que 0,50) , altas foram subescala da subescala SD relacionada ao desconforto subjetivo, ou seja, área de sintomatologia foram com escore total EAS-40 dimensão Somatização (respectivamente 0,83 e 0,80), o que indica que a o escore total da EAS-40 mede constructos parecidos com a dimensão SD .

As correlações entre os escores totais das duas escalas foram de 0,78 para os pacientes e 0,74 para os não pacientes, que ainda podem ser consideradas altas. Porém pode ser visto que a maior correlação da EAS-40 foi com a subescala de Desconforto Subjetivo (0,83) para os pacientes, mostrando uma maior relação com esta do que com o total da EAS-40. Para os não pacientes SD teve uma correlação similar (0,75) ao do resultado total (0,74). Deve ser ressaltado que todas as correlações foram estatisticamente significantes.

As correlações no grupo de pacientes variaram de 0,53 a 0,83 enquanto que no grupo de não pacientes variaram de 0,25 a 0,75. Os resultados entre escore total e das quatro dimensões da EAS-40 com OQ-45.2 e a subescala SD indicaram correlações altas, exceto dimensão Somatização com escore total e SD (OQ-45.2) no grupo de não pacientes em que foram encontradas correlações moderadas, o que pode estar relacionado a uma capacidade dos instrumentos em diferenciar a percepção do estado subjetivo nos dois grupos, o grupo de não pacientes parece diferenciar sintomas psicossomáticos de Desconforto subjetivo.

As correlações entre os escores total do EAS-40 e das suas quatro dimensões com as subescalas IR e SR, nos dois grupos, variaram de 0,25 a 0,65. Os resultados indicaram uma correlação alta entre Somatização com IR e SR nos grupos de pacientes foram 0,60 e 0,63 enquanto no grupo de não pacientes as correlações foram baixas (0,33 e 0,25). Houve um contraste entre as correlações obtidas entre as dimensões Somatização, Obsessividade-compulsividade da EAS-40 e a subescala SR do OQ-45.2, no grupo de pacientes foram obtidas correlações altas enquanto que no grupo de pacientes as correlações foram baixas, pode ser levantar uma hipótese do quanto o sofrimento psíquico dos pacientes são percebidos em diferentes áreas influenciando o modo como estes se relacionam com as pessoas e no desempenho do papel social.

Inventário de Depressão de Beck (BDI-II)

As correlações entre os escores totais e os parciais do Inventário de Depressão de Beck (BDI-II) e o OQ-45.2 são apresentadas na Tabela 15.

Tabela 15. Correlações entre o BDI-II e o OQ-45.2 (N= 189)

	(BDI-II)		
OQ-45	Total	somato-afetiva	cognitiva
Total	0,83*	0,79*	0,76*
SD	0,83*	0,79*	0,77*
IR	0,60*	0,57*	0,55*
SR	0,50*	0,44*	0,50*

* p < 0,001

Foi obtida uma correlação alta e significativa entre os escores totais dos dois instrumentos (0,83), bem como entre o BDI total e SD (0,83). Entre as subescalas IR (0,60) e SR (0,50) e o escore total do BDI foram moderadas. Quanto às dimensões do BDI II, dimensão somato-afetiva e a dimensão psicológica e cognitiva da depressão apresentaram alta correlação com o escore total do OQ (0,79 e 0,76) e com a subescala SD (0,79 e 0,77), indicando a validade do OQ total e SD para avaliar o desconforto subjetivo, entre as quais se incluem sintomas depressivos. Ao comparar os escores do BDI II com as subescalas IR e SR foram obtidas correlações moderadas. Todas as correlações foram significativas ($p \leq 0,001$).

No estudo de Lambert et al. (2004) o BDI também foi utilizado para obter a validade concorrente da subescala SD e o escore total. Foram obtidas as seguintes correlações para SD = 0,63 e escore total = 0,80, que foram mais baixas do que as da amostra brasileira. Na Tabela 16 podem ser observadas as correlações entre os escores do BDI II e OQ-45.2 obtidos com os dois grupos da pesquisa.

Tabela 16. Correlações entre a BDI- II e o OQ-45.2 dos dois grupos (P=22 NP=167)

	BDI-II					
OQ-45.2	Escore total		somato-afetiva		cognitiva	
	P	NP	P	NP	P	NP
Total	0,88*	0,74*	0,80*	0,70*	0,89	0,67*
SD	0,89*	0,75*	0,80*	0,70*	0,80	0,67*
IR	0,82*	0,49*	0,70*	0,47*	0,79*	0,42*
SR	0,46*	0,42*	0,37*	0,34*	0,46*	0,44*

* p < 0,001

P = Paciente; NP = Não paciente

Todas as correlações entre os dois instrumentos foram significantes ($p < 0,001$)

estatisticamente, bem como foram mais altas no grupo de pacientes, indicando que as duas escalas são sensíveis para detectar os problemas emocionais apresentados pelos pacientes. No grupo de pacientes as correlações foram muito altas entre escores total do BDI-II e escores total e das subescalas SD e IR, mas baixa com SR (0,46). Os escores das escalas somato-afetiva e cognitiva da BDI-II com os escores totais do OQ-45.2, SD e IR tiveram correlações foram altas. A correlação obtida foi muito alta entre o escore da dimensão cognitiva da depressão e o escore total do OQ-45.2. Foram obtidas correlações entre moderadas e baixas nos dois grupos, ao comparar os escores do BDI-II com a subescala SR do OQ-45.2, essas variaram de 0,34 a 0,46.

Questionário Geral de Saúde (QSG)

Os coeficientes de correlação para a amostra total com o dos escores do QSG e do OQ-45.2, são mostrados na Tabela 17.

Tabela 17. Correlações entre o QSG e o OQ-45.2 (N= 115)

QSG	OQ-45.2			
	TOTAL	SD	IR	SR
Estresse	0,83*	0,82*	0,71*	0,53*
D. Morte	0,83*	0,81*	0,70*	0,55*
Desconfiança	0,82*	0,80*	0,71*	0,55*
Sono	0,53*	0,51*	0,45*	0,37*
Psicossom.	0,71*	0,72*	0,50*	0,50*
Saúde Geral	0,88*	0,86*	0,73*	0,59*

* $p < 0,01$

As correlações entre as pontuações dos três fatores (estresse, desejo de morte, desconfiança no próprio desempenho) e o escore total do QSG com o escore total do OQ foram muito altas. A correlação entre o fator de distúrbios psicossomáticos e escore total OQ foi alta (0,71). Com o constructo distúrbios do sono a correlação foi moderada com o resultado total (0,53) e baixa com a subescala SR (0,37), o que é esperado, pois avaliam constructos diferentes.

Ao ser relacionados os escores do fator Desconfiança no próprio desempenho, que se refere à consciência de ser capaz de realizar as tarefas diárias de forma satisfatória e os

escores da subescala Desempenho do Papel Social (SR) foi obtida uma correlação moderada 0,55, um indicativo de que avaliam constructos semelhantes. No geral, os resultados obtidos na correlação do QSG com OQ-45.2 que estes instrumentos estariam medindo constructos semelhantes.

Tabela 18. Correlações entre QSG e OQ-45.2 dos grupos de Não pacientes e (N= 95) e Pacientes (N = 20)

QSG	OQ-45.2							
	TOTAL		SD		IR		SR	
	NP	P	NP	P	NP	P	NP	P
Estresse	0,71**	0,87**	0,69**	0,87**	0,56**	0,81**	0,37**	0,52*
D. Morte	0,66**	0,89**	0,63**	0,88**	0,55**	0,78**	0,32**	0,58**
Descon.	0,66**	0,90**	0,62**	0,90**	0,56**	0,77**	0,33**	0,61**
Sono	0,38**	0,74**	0,37**	0,72**	0,27**	0,78**	0,21**	0,49**
Psicos.	0,61**	0,76**	0,63**	0,78**	0,34**	0,59**	0,36**	0,55*
SG	0,77**	0,92**	0,74**	0,91**	0,59**	0,82**	0,40**	0,61**

** p < 0,01; *p<, 0,05
NP = Não pacientes; P = Paciente

Todas as correlações entre os dois instrumentos foram significantes estatisticamente. A correlação entre a Saúde Geral e o escore total do OQ-45.2 foi de 0,72 para os não pacientes e de 0,92 para os pacientes. As correlações obtidas do grupo de pacientes foram sempre maiores do que as dos não pacientes.

As correlações entre as pontuações dos cinco fatores (estresse, desejo de morte, desconfiança no próprio desempenho, sono e psicotismo) e o escore total do QSG com o escore total do OQ foram altas no grupo de não pacientes. As correlações entre os fatores sono e distúrbios psicossomáticos e o escore total OQ-45.2 foram altas (respectivamente 0,74 e 0,76). As correlações entre as pontuações dos três fatores (estresse, desejo de morte e desconfiança no próprio desempenho) e o escore total do OQ-45.2 foram muito altas no grupo de pacientes. Com os dois fatores (sono e psicotismo) as correlações foram altas.

As correlações entre as pontuações dos quatro fatores (estresse, desejo de morte, desconfiança no próprio desempenho e psicotismo) e escore do SD foram altas no grupo de não pacientes. A correlação entre o fator sono e SD do OQ-45.2 foi baixa (0,37) um indicativo de que não avaliam os mesmos constructos.

No grupo de pacientes as correlações entre as pontuações dos três fatores (estresse, desejo de morte e desconfiança no próprio desempenho) e o escore total do QSG foram muito altas no grupo de pacientes. Com os dois fatores (sono e psicotismo) as correlações foram altas. Os dados obtidos indicam que subescala SD, os fatores e escore total SG do QSG avaliam constructos semelhantes.

As correlações entre as pontuações dos cinco fatores (estresse, desejo de morte, desconfiança no próprio desempenho e psicotismo) do QSG com a subescala SR do OQ-45-2 foram baixas (variando de 0,21 a 0,37) no grupo de não pacientes e para o escore total do QSG (SG = 0,40) correlação foi moderada. As correlações foram altas entre o fator desconfiança no próprio desempenho e o escore total do QSG com SR (desempenho do papel social),

De modo geral, os resultados obtidos mostram indicadores de que a versão brasileira do OQ-45.2 é válida e mostra sensibilidade para diferenciar grupos clínicos de não-clínicos.

CONSIDERAÇÕES FINAIS

O objetivo principal deste estudo foi obter evidências de precisão e validade da versão brasileira do *Outcome Questionnaire* (*OQ-45.2*). A precisão foi estabelecida pelo método teste-reteste e pela consistência interna por meio do alfa de Cronbach da escala total e das suas subescalas. A validade concorrente foi investigada empregando como critérios os seguintes instrumentos: Escala de Avaliação de Sintomas (EAS-40), Inventário de Depressão de Beck (BDI-II) e Questionário Geral de Saúde (QSG). Foi estabelecida a validade por meio de grupo de contrastes entre pacientes e não pacientes para a pontuação total da escala e das suas subescalas. Também foram calculadas as intercorrelações entre o escore total e as subescalas e verificadas as diferenças entre as amostras clínica e não clínica para cada resultado do instrumento.

Na versão brasileira do OQ-45.2 foram obtidos coeficientes correlação de Pearson para o reteste com valores maiores do que os relatados no estudo com a versão original da escala (Lambert et al., 2004) e da versão holandesa (Jong et al., 2007). Para o escore total, a subescala SD e IR, exceto para a subescala SR que obteve menor correlação (0,58) na comparação com os dois estudos estrangeiros americano e holandês.

Quanto à validade, no caso de uma escala considerada multidimensional, que avalia diferentes aspectos do funcionamento, as correlações entre as subescalas, não devem ser muito altas, uma vez que cada subescala deve avaliar conceitos que não são medidos nas outras subescalas. No caso versão brasileira do OQ-45.2, a subescala SD que mede sintomas de desconforto subjetivo teve excelente correlação com escore total nos dois grupos estudados. Foram obtidas correlações altas ao se relacionar o escore total com as subescalas IR, que avalia os relacionamentos interpessoais e SR, referente ao desempenho do Papel Social, entre as subescalas SD e IR as correlações entre as subescalas, que avaliam desconforto subjetivo e relações interpessoais foram moderadas. Os resultados indicam que a subescala SD apresenta correlações muito altas com a escala total nos dois grupos estudados e alta entre escore total IR e SR. As correlações moderadas entre SD e IR e IR com SR nos dois grupos.

As diferenças entre as médias obtidas nos dois grupos mostraram que o Q-45. 2 é capaz de diferenciar grupos de pacientes de não pacientes, bem como foram obtidas médias maiores para os paciente, o que é uma confirmação da sua validade de 82

constructo. Enquanto na amostra americana não foram identificadas diferenças quanto ao gênero, isto não ocorreu nas amostras italiana, holandesa e brasileira.

Este estudo representa uma contribuição importante para uma futura utilização e publicação de um manual brasileiro do OQ-45.2, uma vez que mostrou que a adaptação realizada por Carvalho e Rocha (2009) apresenta dados satisfatórios de validade e precisão, bem com consegue diferenciar entre os pacientes e os não pacientes. Para complementar a presente pesquisa ainda é necessário realizar um estudo para verificar a validade fatorial da escala e determinar se os itens traduzidos irão se distribuir em fatores do mesmo modo que os estudos originais, bem como de acordo com as subescalas propostas pelos autores.

Também é fundamental a realização de um estudo normativo com uma amostra representativa para a determinação dos pontos de corte e parâmetros para a escala. Também são necessários estudos com diferentes amostras, de modo que as tomada de decisões para o tratamento possa ser baseadas em normas para a população, evitando decisões inadequadas para a condução da psicoterapia.

Uma das limitações deste estudo está relacionada ao tamanho da amostra de pacientes (N=59), mas mesmo assim esta mostrou diferenças em relação ao grupo de controle.

Outro aspecto importante que necessita ser investigado refere-se à sensibilidade da escala para avaliar as mudanças ocorridas na psicoterapia e que constitui o objetivo principal para o qual esta escala foi criada. Esse é um desafio a ser enfrentado, principalmente se se considerar as diversas variáveis que podem influir sobre os resultados de uma psicoterapia, incluindo-se as diversas abordagens teóricas, as diferenças entre os psicoterapeutas e entre os próprios clientes.

REFERÊNCIAS

Abanishe, D. R. (2008). *Cros-cultural comparison of the Outcome Questionnaire*. Tese de doutorado em filosofia. Universidade de Oklahoma. Norman: Oklahoma. *ProQuest: UMI 3304221*.

Alves, H. (2001*). Psicoterapia breve operacionalizada-PBO com adolescentes da rede pública escolar da cidade de Santos-SP: Uma contribuição à psicologia Clínica preventiva.* Tese de Doutorado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Anastasi, A. (1977). *Testes Psicológicos*. São Paulo: EPU.

Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica*. Porto Alegre: ArtMed.

Ankuta, G. Y. & Ables, N. (1993). Client satisfaction, clinical significance, and meaningful change in psychotherapy. *Professional Psychology: Research and Practice, 24* (1), 70-74.

Associação Americana de Psicologia (2013). *Recognition of psychoterapy effectiveness. Psychotherapy, 50* (1), 102-109.

Associação Brasileira de Empresas de Pesquisa (2010). Critério de Classificação Econômica Brasil. Disponível em: <http://www.abep.org/novo/Content.aspx? ContentID=301>. Acesso: 10 out. 2010.

Azevedo, M. (2006). *Psicoterapia para o povo: Avaliação de resultados. Psicologia para America Latina, 8*. Disponível em: <http://pepsic.bvsalud.org/scielo.php? script=sci_arttext&pid=S1870-350X2006000400013&lng=pt&nrm=iso>. Acesso: 24 mar. 2011.

Azevedo, M. (2009). Avaliação dos resultados de intervenções psicoterápicas breves na comunidade de Bauru. *Revista Ciência em Extensão, 2* (1). Disponível em: <http://ojs.unesp.br/index.php/revista_proex/article/view/184/98> Acesso: 24 mar. 2011.

Ballesteros, S. T., Vitriol, G. V., Florenzano, U. R., Vacarezza, L. A. & Calderón, K. A. (2007). Mujeres con depresión severa: Relación entre trauma infantil y gravedad de sintomas clínicos. *Revista Chilena de Neuro-Psiquiatri, 45* (4), 288-295.

Beckstead, D. J., Hatch, A. L., Lambert, M. J., Eggett, D. L., Goates, M. K. & Vermeersch, D. A. (2003). Clinical significance of the Outcome Questionnaire (OQ-45.2). *The Behavior Analyst Today, 4* (1), 79-90.

Benetti, S. P. C. & Cunha, T. R. S. (2008). Abandono de tratamento psicoterápico: Implicações para a prática clínica. *Arquivos Brasileiros de Psicologia, 60* (*2*), 48-59.

Beretvas, S. N. Kearney, L. K. & Barón, A. (2003). *A shortened form of the Outcome Questionnaire: A validation of scores across ethnic groups*. (Relatório do grupo em pesquisa de aconselhamento e serviços de Psicologia no Ensino Superior). Universidade do

Texas: Austin. Disponível em: <http://www.utexas.edu/ student/cmhc/research/rescon.html>. Acesso: 03 mar. 2012.

Biaggio, A. M. B. & Natalicio, L. (1979). *Manual para Inventário de Ansiedade Traço-Estado (IDATE)*. Rio de janeiro: CEPA. 84.

Bortolini, M., Pureza, J. R., Andretta, I. & Oliveira, M. S. (2011). Perfil de pacientes atendidos através da terapia cognitivo-comportamental em uma clínica-escola. *Contextos Clínicos, 4* (2), 132-138.

Bunchaft, G. & Cavas, C. T. (2002). *Sob Medida: um guia sobre a elaboração de medidas do comportamento e suas aplicações*. São Paulo: Vetor.

Callahan, J. L. & Hynan, M. T. (2005). Models of psychotherapy outcome: Are they applicable in training clinics? *Psychological Services, 2* (1), 65-69.

Campezatto, P. M. & Nunes, M. L. T. (2007). Atendimento em clínicas-escola de psicologia da região metropolitana de Porto Alegre. *Estudos de Psicologia, 24* (3), 363-374.

Carvalho, L. F. & Rocha, G. M.A. (2009). Tradução e adaptação cultural do Outcome Questionnaire (OQ-45). *Psico-USF, 14* (3), 309-316.

Cerezetti, C. R. N. (2002). *O paciente de transplante hepático: Um estudo psicológico sobre a força resiliente do hepatopata grave*. Dissertação de Mestrado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Chilelli, K. B. & Enéas, M. L. E. (2000). Desistência em psicoterapia breve: Pesquisa documental e da opinião do paciente. *Boletim de Iniciação Científica de Psicologia, 1* (1), 47-52.

Cisneros, E. P. (2010). *Client/therapist feedback and the role of the aliance on psychotherapy outcomes.* Dissertação de mestrado. Loma Linda University School: California. Disponível em: <http://proquest.umi.com/pqdweb?did= 2169244901&sid=1&Fmt=2&cl = 18113 & RQT = 309 & VName = PQD Documento ProQuest ID: 2169244901>. Acesso: 30 mar. 2013.

Conselho Federal de Psicologia (2000). *Especifica e qualifica a psicoterapia como prática do psicólogo*. (Resolução Nº 010/2000). Brasília: CFP. Disponível em: <http://site.cfp.org.br/wp-content/uploads/2000/12/resolucao2000_10.pdf>. Acesso: 21 abr. 2013

Conselho Federal de Psicologia (2003). *Define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos e revoga a Resolução CFP n° 025/2001.* (Resolução Nº 002/2003). Brasília: CFP. Disponível em: <www.pol.org.br/legislacao/ doc/resolucao2003_2.doc>. Acesso: 30 mar. 2013.

Conselho Federal de Psicologia (2009). *Sistematização do Seminário Nacional do Ano da Psicoterapia.* (2009 Ano da Psicoterapia). Brasília: CFP. Disponível em:

<http://site.cfp.org.br/wp-content/uploads/2009/05/Ano-da-Psicoterapia-Textos-geradores.pdf>. Acesso: 20 abr. 2013.

Conselho Federal de Psicologia (2011). *Ano da Avaliação Psicológica.* (Textos Geradores). Brasília: CFP. Disponível em: <http://site.cfp.org.br/wp-content/ uploads/2013/04/anodaavaliacaopsicologica_prop8.pdf>. Acesso: 20 abr. 2013. 85.

Correa, D. J. , Florenzano, U. R. Rojas, H. P., Labra, J. J. F., Río, D. V. del, Pastén, J. A. (2006). El uso del cuestionario OQ-45.2 como indicador de psicopatologia y de mejoría em pacientes psiquiátricos hospitalizados. *Revista Chilena Neuro-psiquiatria, 44*(4), 258-262. Disponível em: <http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-92272006000400004&lng=es&nrm=iso>. Acesso em 30.4.2012.

Dantas, M. S. (2007). *Psicoterapia breve de problemas conjugais: Informatização de registros na clínica institucional*. Dissertação de Mestrado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

De La Parra, G. & Bergen, V. (2002). OQ 45-2 Cuestionario para evaluación de resultados y evolución en psicoterapia: adaptación, validación, e indicaciones para su aplicación e interpretación. *Revista de Terapia*, *20*, 161-176.

Dell´Aquila, D. C., Castro, E. G. & Mosca, A. S. (2009). Obesidad y adolescencia: exploración de aspectos relacionales y emocionales. *Terapia Psicológica, 27* (1), 143-149.

Del Prette, G. (2006). *Terapia analítico-comportamental infantil: relações entre o brincar e comportamentos da terapeuta e da criança*. Dissertação de Mestrado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Del Prette, G. (2011). *Objetivos analítico-comportamentais e estratégias de intervenção nas interações com a criança em sessões de duas renomadas terapeutas infantis*. Tese de Doutorado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Doerfler, L. A., Addis, M. E. & Moran, P. W. (2002). Evaluating mental health outcomes in an inpatient setting: Convergent and divergent validity of the OQ-45 and BASIS-32. *The Journal of Behavioral Health Services & Researc, 29* (4), 394-403.

Draper, M. R., Jennings, J. & Barón, A. (2003). Factor analysis and concurrent validity of a University Counseling Center presenting Problems Checklist. Disponível em: <http://www.utexas.edu/student/cmhc/research/rescon.html>. Acesso: 30 mar. 2013.

Duran, S. M. G. T (1997). *O atendimento psicoterapêutico em grupo aos usuários de uma Unidade Básica de Saúde pelo Método Corporal Peth Sándor: Uma interpretação na perspectiva da psicologia analítica de C. G. Jung*. Dissertação de Mestrado. Instituto de psicologia da Universidade de São Paulo, São Paulo.

Enéas, M. L. E. (2003). Mudança psíquica avaliada pela Escala Rutgers de Progresso em

Psicoterapia. *Avaliação Psicologica*. *2* (1). Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712003000100006&lng=pt&nrm=iso>. Acesso: 30 mar. 2013.

Enéas, M. L. E. (2008). Pesquisa em psicoterapias: Seções Especiais de Periódico (1981-1994). *Psicologia: Teoria e Prática, 24* (1),111-116.

Enéas, M. L. E., Faleiros, J. C. & Sá, A. C. A. (2000). Uso de psicoterapias breves em clínica escola: caracterização dos processos com adultos. *Psicologia Teoria e Prática. 2* (2), 9-30.

Florenzano, R., Artero, M., Gonzalez, I. & Martinez, X., Ortega, P.(2005). Evaluación del impacto de las intervenciones psicosociales em el nível hospitalário de uma Red de Atención de Salud Mental: Aplicación sistemática del OQ-45.2 de Lambert. *Alcmeon Revista Argentina de Clínica Neuropsiquiátrica, 12* (2), 46. Disponível em: <www.alcmeon.com.ar>. Acesso: 20 mar. 2013.

Freire, E. S. (2006). *Desenvolvimento de um instrumento de avaliação de resultados em psicoterapia baseado na teoria da mudança terapêutica de Carl Rogers*. Tese de Doutorado. Universidade Federal do Rio Grande do Sul, Porto Alegre.

Freud, S. (1912/1969). Recomendações aos médicos que exercem a Psicanálise. Edição standard das obras completas de Sigmund Freud. Vol. XII, (pp. 149-159). Rio de Janeiro: Imago.

Gatti, A. L. (1994). *Alterações no Rorschach após psicoterapia psicodinâmica breve.* Dissertação de Mestrado. Pontifícia Universidade Católica de Campinas, Campinas.

Gatti, A. L. & Jonas, A. L.(2007). Caracterização do atendimento psicoterápico a adultos em clínica-escola no ano de 2005. *Integração, 49*, 89-93.

Gebara, A. C. (2009*). Pesquisa sobre a eficácia terapêutica da interpretação teorizada na psicoterapia breve operacionalizada*. Tese de Doutorado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Giovannetti, R. M. & Sant´Anna, P. A. (2005). Estratégias de psicodiagnóstico interventivo e apoio em crises adaptativas por meio do Jogo de Areia e da EDAO. *Psicologia: Reflexão e Crítica, 18* (3), 402-407.

Gorenstein, C., Pang, W. Y., Argimon, I. L. & Werlang, B. S. G. (2011) *Manual do Inventário de Depressão de Beck - BDI-II.* São Paulo: Casa do Psicólogo.

Hansen, N. B., Lambert, M. J. & Forman, E. M. (2002). The Psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science Practice, 9* (3), 329-343.

Harmon, C., Hawkins, E.J., Lambert, M. J., Slade, K. & Whipple, J.S. (2005). Improving outcomes for poorly responding clients: The use of Clinical Support Tools and feedback to clients. *A Journal of Clinical Psychology, 61* (2), 175-185.

Haug, S., Puschner, B., Lambert, M.J. & Kordy, H. (2004). Veränderungsmessung in der Psychotherapie mit dem Ergebnisfragebogen (EB-45). Zeitschrift für Dijferentielle und Diagnostische. *Psychologie, 25*, 141-151.

Hilsenroth, M. J. (2013). Introduction to the 50th anniversary special issue on psychotherapy outcome: A return to the beginning. *Psychotherapy, 50* (1), 1-2.

Honda, G. C. (2010). *Avaliação de mudança em pacientes de clínica-escola*. Dissertação de Mestrado. Pontífice Universidade Católica de Campinas, Campinas.

Horvath, A. O. & Greenberg, L. S. (1994). The Working Alliance.Theory, Research and Practice. New York, John Wiley & Sons.

Howard, H. I., Kopta, S. M. & Krause, M. S. (1986). The dose-effect relationship in psychotherapy. *American Psychological, 41*, 159-164.

Jacobson, N. S., Follette, W. C. & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*, 336-352.

Jacobson, N. S. & Truax, P.(1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consuling and Clinical Psychology, 59*, 12-19.

Jong, K., Nugter, M. A., Polak, M. G., Wagenborg, J. E. A., Spinhoven, P. & Heiser, W. J. (2007). The Outcome Questionnaire (OQ-45) in a Dutch population: A cross-cultural validation. *Clinical Psychology and Psychotherap, 14* (4), 288-301.

Jung, S. I. (2007). *Avaliação de resultados da psicoterapia psicanalítica em um serviço de atendimento em Porto Alegre*. Dissertação de Mestrado. Programa de Pós-Graduação em Ciências Médicas. Universidade Federal do Rio Grande do Sul

Jung, S. I., Fillipon, A. P. M., Nunes, M. L. T. & Eizirik, C. L. (2006). História recente e perspectiva atuais da pesquisa de resultados em psicoterapia psicanalítica de longa duração. Revista de Psiquiatria Rio Grande do Sul, *28* (3), 298-312.

Kadera, S. W., Lambert, M. J. & Andrews, A. A. (1996). How Much Therapy Is Really Enough? A Session-by-Session Analysis of the Psychotherapy Dose-Effect relationship. *Journal Psychotherapy Practice Research, 5* (2), 132-151.

Kearney, L. K., Draper, M. & Barón, A. (2003). Couseling utilization by ethnic minority students. *The Counseling & Mental Health Center,* 1-23. Disponivel em: <www.utexas.edu/Student/research/rescon.html>. Acesso: 20 mar. 2013.

Kim, J., Nugter, M. A., Polak, M. G., Wagenborg, J. E. A., Spinhoven, P. & Heiser, W. J. (2007). The Outcome Questionnaire (OQ-45) in a Dutch population: A cross-cultural validation. *Clinical Psychology & Psychotherapy, 14* (4), 288-301.

Laham, C. F., Satake, P. S., Chiba, T., Benute, G. R. G. & Lucia, M. C. S. (2009). O Uso da escala Diagnóstica Adaptativa Operacionalizada (EDAO) no domicílio: Um estudo descritivo. *Psicologia Hospitalar, 7* (1), 68-87.

Laloni, D. T. (2001). *Escala de Avaliação de Sintomas-90-R SCL-90-R: Adaptação, precisão e validade*. Tese de Doutorado. Pontifícia Universidade Católica de Campinas, Campinas.

Laloni, D. T. (2008). Escala de Avaliação de Sintomas -90-R. SCL-90-R: Adaptação, Precisão e Validade. In: A. P. Noronha, et al. (Org.), *Actas do XIII Conferência Internacional de Avaliação Psicológica: Formas e Contextos.* Braga: Psiquilibrios Edições.

Lambert, M. J. (1983). Introduction to assessment of psychotherapy outcome: Historical perspesctive and current issues. In: M. J. Lambert, E. R. Christensen & S. S. DeJulio (Eds.), *The assessment of psychotherapy outcome,* (pp. 3-32). New York: John Wiley and Sons.

Lambert, M.J. (2007). Presidential address: what we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research. 17* (1), 1-14. New York: Routledge.

Lambert, M. J. (2013). Outcome in Psychotherapy: The past and important advances. *Psychotherapy, 50* (1), 42-51.

Lambert, M. J., Burlingame, G. N., Umphress, V., Hanse, N.B., Vermeersh, D. A., Clouse, G. C. & Yanchar, S. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy*, *3*, 249-258. 89

Lambert, M. J. & Finch, A. E. (1999). The Outcome Questionnaire. M. E. Maruish (Ed.). The use of psychological testing for treatment planning and outcomes assessment. (pp. 831-869). New Jersey: Lawrence Erlbaum Associates, Publishers.

Lambert, M. J., Hansen, N.B. & Finch, A.E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology, 69*, 159-172.

Lambert, M. J., & Hill, C. (1994). Assessing psychotherapy outcomes and processes. In A.E. Bergin and S. L. Garfield (eds.). *Handbook of psychotherapy and behavior change* (*4*), 72-113). New York: John Wiley and Sons.

Lambert, M. J., Morton, J. J., Hartfield, D, Harmon, C., Hamilton, S. Reid, R. C., Shimokawa, K., Christopherson, C. & Burlingame, G. M. (2004). *Administration and Scoring Manual for the OQ-45.2 Outcome Questionnaire.* Salt Lake City: American Professional Credentialing Services.

Lambert, M. J., Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change.*

(*5*) 805-821. New York: Wiley.

Lambert, J. M., Smart, D. W., Campbell, M. P., Hawkins, E. J., Harmon, C. & Slade, K. L. (2006). Psychotherapy outcome, as measured by the OQ-45, in African American, Asia/Pacific Islander, Latino/a, and Native American clients compared with matched caucasian clients. *Journal of College Student Psychotherapy*. *20* (4), 17-29.

Lara, C., Cruz, C., Vacarezza, A., Florenzano U. R. & Trapp, A. (2008). Análisis comparativo de dos instrumentos de evaluación clínica: OQ45 e *InterRAI* – Salud Mental. *Revista Chilena de NeuroPsiquiatria, 46* (3),192-198.

Lo Coco, G. L., Chiappelli, M., Bensi, L., Gullo, S., Prestano, C., & Lambert, M. J. (2008). The factorial structure of the Outcome Questionnaire-45: A study with an Italian sample. *Clinical Psychology and Psychotherapy, 15*, 418-423.

Luborsky L, McLelan, A.T., Woody G.E., O'Brien C.P. & Auerback A. (1985). Therapist success and it determinants. *Arch Gen Psychiatry, 42*, 602-611.

Lunnen, C. & Ogles, B. M. (1998). A multi-perspective, muti-variable evaluation of effect change. *Journal of Consulting and clinical Psychology*, *66*, 400-410.

Machado, P. P. P., Klein, J. M., Farate, C. (2005). Monitorização dos resultados terapêuticos no Contexto de tratamento das Toxicodependências. *Psicologia: Teoria e Investigação e Prática*, *1* (1), 19-29.

Marcolino, J. A. M. & Iacoponi, E. (2001). Escala de Aliança Psicoterápica da California na versão paciente. *Revista Brasileira de Psiquiatria, 23* (2), 88-95.

Martínez, C., Alonso, C., Castro, P., Álvarez, P., Oviedo, C. & Ellwanger, J. (2004). Efectividad de la técnica de intervención em crisis (IC) em um servicio de urgência de um hospital general. *Terapia Psicológica*, *22* (2), 177-184.

Massei, A. C. (2009). *Propriedades Psicométricas do Questionário de Relacionamento Central (CRQ) em portadores de HIV/AIDS*. Dissertação de Mestrado. Pontifícia Universidade Católica de Campinas, Campinas.

Melnik, T. (2006). *A efetividade da psicoterapia no tratamento da disfunção erétil: revisão sistemática e metanálise*. Tese de Doutorado. Instituto de psicologia da Universidade de São Paulo, São Paulo.

Meyer, S. B., Oshiro, C., Donadone, J. C., Mayer, R. C. F. & Starling, R. (2008). Subsídios da obra "Comportamento Verbal" de B. F. Skinner para a terapia analítico-comportamental. *Revista Brasileira de Terapia Comportamental e Cognitiva*. *X* (1), 105-118. Disponível em: <http://pepsic.bvsalud.org/pdf/rbtcc/v10n1/v10n1a09.pdf>. Acesso: 30 mar. 2013.

Morgan, R., Luborsky, L., Crits-Christoph P., Curtis & H. Solomon, J. (1982). Predicting the outcomes of psychotherapy by the Penn Helping Alliance Rating Method. *Arch Gen*

Psychiatr, 39 (4), 397-402.

Mueller, R. M., Lambert, M.J. & Burlingame, G. M. (1998). Construct validity of the Outcome Questionnaire: A confirmatory factor analysis. *Journal of Personality Assessment, 70* (2), 248-262.

Nelson, P. L. (2010). Change trajectories and early warning system to identify youth at risk for negative psychotherapy outcome. Dissertação de Mestrado. Departamento de Psicologia. Brigham Young University. Idaho. Disponivel em: <http://proquest.umi.com/pqdweb?did=2192433681&sid=1&Fmt=2&clientId=79356&RQT=309&VName=PQD>. Acesso: 05 jan. 2013

Neno, S. (2005). Tratamento padronizado: condicionantes históricos, status contemporâneo e (in) compatibilidade com a terapia analítico-comportamental. Tese de Doutorado. Programa de Pós-Graduação em teoria e Pesquisa do Comportamento. Universidade Federal do Pará: Belém.

Okiishi, J. C., Lambert, M. J., Eggett, D., Nielsen, L. & Dayton, D., D. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their clients´psychotherapy outcome. *Journal of Clinical* Psychology, *62* (9), 1167-1172.

Orlinsky, D. & Horward K. (1986). The psychological interior of psychotherapy: exploration with the therapy session reports. In: L, Grenberg & W, Pinso ., (Eds.).*The* 91 *psychotherapeutic process: A research handbook,* 477-500. Nova York: Guilford Press.

Oshiro, C. K. B. C.(2011). *Delineamento Experimental de Caso Único: A psicoterapia analítica funcional em dois clientes difíceis*. Tese de Doutorado. Instituto de Psicologia. Universidade de São Paulo, São Paulo.

Pasquali, L.(2003) *Psicometria: Teoria dos testes na Psicologia e na Educação*. Rio de Janei ro: Vozes.

Paquali, L., Gouveia, V. V., Andriola, W. B., Miranda, F. J. & Ramos, A. L. M. (2011). *QSG Questionário de Saúde Geral de Goldberg*. 1 reimp. São Paulo: Casa do Psicólogo.

Peres, R.S. (2009). Aliança terapêutica em psicoterapia de orientação psicanalítica: aspectos teóricos e manejo clínico. *Estudos de Psicologia. 26* (3)*, 383-389.*

Potter, A. R., Baker, M. T., Sanders, C. S. & Peterson, A.L. (2009). Combat stress reactions during military deployments: Evaluation of the effectiveness of combat stress control treatment. *Journal of Mental Health Counseling, 31* (2), 137-148.

Prado, O. Z. & Meyer, S. B.(2006). Avaliação da relação terapêutica na terapia assíncrona via internet. *Psicologia em Estudo*, 11 (2), 247-257. Disponível em: <www.scielo.br/pdf/pe/v11n2/v11n2a02.pdf>. Acesso: 20 mar. 2013.

Ramos, M. A. F. (2008). *Análise das características psicométricas da Versão Portuguesa do*

Working Alliance Inventory – Short Revised. Dissertação de Mestrado. Universidade do Minho, Minho.

Reese, R. J., Norsworth, L. A. & Rowlands, S. R. (2009). Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy Theory, Research, Practice, Training. 46* (4), 418-431.

Rocha, G. M. A. (2002). *Escala Diagnostica Adaptativa Operacionalizada Redefinida: precisão e validade com pessoas idosas*. Dissertação de Mestrado. Pontifícia Universidade Católica de Campinas. Campinas.

Rodríguez, M. A. von B. (2000). *Investigación Empírica em Psicoterapia: Validación del Cuestionario de Resultados Terapêuticos OQ-45.2*. Tesis para optar al grado de Licenciado em Psicologia. Universidad Nacional Andres Bello. Santiago do Chile.

Sanches, F. M. (2009). *Questionário de Relacionamento Central (CRQ): Evidências de Validade em pacientes cardíacos*. Dissertação de Mestrado. Pontifícia Universidade Católica de Campinas, Campinas.

Santeiro, T. V. (2005). *Psicoterapias breves psicodinâmicas: Produção cientifica em periódicos nacionais e estrangeiros (1980/2002).* Tese de Doutorado. Pontifícia Universidade Católica de Campinas, Campinas.

Serralta, F. B., Nunes, M. L. T., & Eizirik, C. L. (2007). Elaboração da versão em português do Psychotherapy Process Q-Set. *Revista Brasileira de Psiquiatria*, 29 (1), 44-55.

Shimokawa, K. (2010). *A Patient-Focused Psychotherapy Quality Assurance System: Meta-Analytic and Multilevel Analytic Review*. Tese de Doutorado. Departamento de Psicologia. Brigham Young University. Idaho. Disponível em: <http://proquest.umi.com/pqdweb?did=2192423731&sid=1&Fmt=2&clientId=18113&RQT=309&VName=PQD>. Acesso: 01 mar. 2012.

Simon, R. (1989). *Psicologia Clínica Preventiva: novos fundamentos.* São Paulo: EPU.

Simon, R. (2010). *Psicoterapia psicanalítica: Concepção original. Teoria técnica pesquisa ilustrações clínicas.* São Paulo: Casa do Psicólogo.

Storch, A. N. (2010). *Analyzing treatment effectiveness in deaf clients who use a signing versus nonsigning therapist*. Dissertação de Mestrado. Walden University. Disponível em: <http://proquest.umi.com/pqdweb?did=2201873611&sid=1&Fmt=2&cl%20ientId=18113&RQT=309&VName=PQD&cfc=1>. Acesso: 20 mar. 2013.

Strupp, H. H. (2013). The outcome problem in psychotherapy: a rejoinder. *Psychotherapy, 50* (1), 15.

Strupp, H.H., Horowitz, L.M. & Lambert, M. (Eds.) (1997). *Measuring patient changes in mood, anxiety, and personality disorders: toward a core battery*. Washington, DC: American

Psychological Association.

Taccola, P. A. (2007). *A psicoterapia analítico-funcional e relato de sentimentos: um estudo de caso quase experimental*. Dissertação de Mestrado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Umpress, V. J., Lambert, M. J., Smart, D. W., Barlow, S. H., & Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. *Journal of Psychoeducational Assessment, 15*, 40-55.

Urbina, S. (2007). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artmed.

Van Kolck, L. (1974). *Técnicas de exame psicológico e suas aplicações no Brasil*. Petrópolis: Ed. Vozes.

Worthen, V. E. & Lambert, M. J. (2007). Outcome oriented supervision: Advantages of adding systematic client tracking to supportive consultations. *Counseling and Psychotherapy Research. 7* (1):48-53.

Yano, Y. (2003). *Tratamento padronizado e individualizado no Transtorno do Pânico*. Tese de Doutorado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Yoshida, E. M. P. (1989). *Estudo de critérios psicodiagnósticos para indicação de psicoterapia breve*. Tese de Doutorado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Yoshida, E.M.P. (1999). EDAO-R: precisão e validade. *Mudanças*. 7 (11), 189-213.

Yoshida, E. M. P. (2000). Toronto Alexithimia Scale-TAS: Precisão e Validade da Versão em português. *Psicologia: Teoria e Prática. 2 (1), 59-74.*

Yoshida, E.M.P. (2005). Recepção, acolhimento, triagem e pesquisa na clínica-escola. Em Melo-Silva, L.L., Santos, M.A. & Simon, C.P. (Orgs.). *Formação em psicologia: serviços-escola em debate*. (pp. 271-280). São Paulo: Vetor.

Yoshida, E. M. P. (2008). Significância clínica de mudança em processo de psicoterapia psicodinâmica breve. *Paidéia, 18* (40), 305-316.

Yoshida, E. M. P. (2012). Psicoterapias Psicodinâmicas. Em M. E. N. Lipp, & E. M. P. Yoshida, (Org.). *Psicoterapias breves nos diferentes estágios evolutivos* (pp. 3-17). São Paulo: Casa do Psicólogo.

Yoshida, E. M. P., Elyseu Jr, S., Silva, F. R. C. S, Finotelli Jr, I., Sanches, F. M., Penteado, E. F., Massei, A. C.& Enéas, M. L. E. (2009). Psicoterapia psicodinâmica breve: estratégia terapêutica e mudança no padrão de relacionamento conflituoso. *Psico-Usf, 14* (3), 275-285.

Yoshida, E. M. P. & Rocha, G. M.A. (2007). Avaliação em psicoterapia psicodinâmica. J. C. Alchieri, (org.). *Avaliação Psicológica: perspectivas e contextos*. (p. 237-288). São Paulo:

Vetor.

Yoshida E. M. P., Santeiro T. V., Santeiro, F. R. M. e Rocha, G. M. A.R. (2005). Psicoterapias breves dinâmicas: características da produção nacional e estrangeira (1980/2003). *Psico-Usf, 10* (1), 51-59.

Yoshida, E.M.P. & Silva, F.R. C. S. (2007). Escala de Avaliação de Sintomas-40 (EAS-40): validade e precisão em amostra não-clinica. *Revista Semestral da associação Brasileira de Psicologia Escolar e Educacional (ABRAPEE). 11* (1), 89-99.

Younes, J. A. (2011). *Psicoterapia breve operacionalizada na crise adaptativa por perda: um estudo exploratório*. Dissertação de Mestrado. Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Zamignani, D. R. (2008). *O desenvolvimento de um sistema multidimensional para a categorização de comportamentos na interação terapêutica.* Tese de Doutorado. Instituto de Psicologia da Universidade de São Paulo. São Paulo.

Zane, N., Hall, G. N., Sue, S., Young, K., & Nunez, J. (2004). Research on psychotherapy with culturally diverse populations. In M. Lambert (Ed.), *Bergin and Garfield's* 94 *handbook of psychotherapy and behavior change,* 5, 767-804. New York: Wiley & Sons.

ANEXO A: PACIENTES

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO AO SUJEITO DE PESQUISA

Gostaria de convidá-lo a participar do projeto de pesquisa "**ESTUDO NORMATIVO DA ESCALA DE AVALIAÇÃO DE RESULTADOS DE PSICOTERAPIA OQ45-2, NA VERSÃO EM PORTUGUÊS**" que se propõe a estabelecer dados normativos para a população de adultos, de 18 anos ou mais, a partir da Escala de Avaliação de Resultados de Psicoterapia (*Outcome Questionnaire-OQ45-2).* Os dados para o estudo serão coletados através do preenchimento de um questionário de dados sobre os participantes, a Escala de Avaliação de Sintomas (SCL-90, Inventário de Beck de Depressão (BDI), o Questionário de Saúde Geral de Goldberg, e a Escala de Avaliação de Resultados de Psicoterapia (OQ45-2). Os instrumentos de avaliação serão aplicados pelo Pesquisador Responsável e auxiliares de pesquisas, tanto os instrumentos de coleta de dados quanto o contato interpessoal oferecem riscos mínimos aos participantes, uma vez que os são questionários de autorrelato e versam sobre a percepção que cada indivíduo tem acerca do seu estado psicológico.

Em qualquer etapa do estudo você terá acesso ao Pesquisador Responsável para o esclarecimento de eventuais dúvidas (no endereço abaixo), e terá o direito de retirar a permissão para participar do estudo a qualquer momento, sem qualquer penalidade ou prejuízo. As informações coletadas serão analisadas em conjunto com a de outros participantes e será garantido o sigilo, a privacidade e a confidencialidade das questões respondidas, sendo resguardado o nome dos participantes (apenas o Pesquisador Responsável terá acesso a essa informação), bem como a identificação do local da coleta de dados.

Caso você tenha alguma consideração ou dúvida sobre os aspectos éticos da pesquisa, poderá entrar em contato com o ***Comitê de Ética em Pesquisa da Universidade Guarulhos*** – Praça Tereza Cristina nº Centro Guarulhos.

Desde já agradecemos a sua colaboração.

Declaro que li e entendi os objetivos deste estudo, e que as dúvidas que tive foram esclarecidas pelo Pesquisador Responsável. Autorizo a publicação dos dados.

Ficou claro também que a minha participação é isenta de despesas e que tenho garantia de acesso a tratamento ou orientação quando necessário. Estou ciente que a participação é voluntária, e que, a qualquer momento tenho o direito de obter outros esclarecimentos sobre a pesquisa e de retirar a permissão para participar da mesma, sem qualquer penalidade ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido, ou à continuidade do meu tratamento na Instituição.

Assinatura do Paciente: _____

RG: _____

Sonia Maria da Silva
Psicóloga CRP 06/13818
Pesquisadora Responsável
End. Rua Dr. Nilo Peçanha, 47
CEP 07011-040
Centro - Guarulhos
Centro – Guarulhos

ANEXO B: Não Pacientes

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO ao Sujeito de Pesquisa

Gostaria de convidá-lo a participar do projeto de pesquisa "**ESTUDO NORMATIVO DA ESCALA DE AVALIAÇÃO DE RESULTADOS DE PSICOTERAPIA OQ45-2, NA VERSÃO EM PORTUGUÊS**" que se propõe a estabelecer dados normativos para a população de adultos, de 18 anos ou mais, a partir da Escala de Avaliação de Resultados de Psicoterapia (*Outcome Questionnaire-OQ45-2)*. Os dados para o estudo serão coletados através do preenchimento de um questionário de dados sobre os participantes, a Escala de Avaliação de Sintomas (EAS-40, o Questionário de Saúde geral de Goldberg, e a Escala de Avaliação de Resultados de Psicoterapia (OQ45-2). Os instrumentos de avaliação serão aplicados pelo Pesquisador Responsável e auxiliares de pesquisas, tanto os instrumentos de coleta de dados quanto o contato interpessoal oferecem riscos mínimos aos participantes, uma vez que os são questionários de autorrelato e versam sobre a percepção que cada indivíduo tem acerca do seu estado psicológico.

Em qualquer etapa do estudo você terá acesso ao Pesquisador Responsável para o esclarecimento de eventuais dúvidas (no endereço abaixo), e terá o direito de retirar a permissão para participar do estudo a qualquer momento, sem qualquer penalidade ou prejuízo. As informações coletadas serão analisadas em conjunto com a de outros participantes e será garantido o sigilo, a privacidade e a confidencialidade das questões respondidas, sendo resguardado o nome dos participantes (apenas o Pesquisador Responsável terá acesso a essa informação), bem como a identificação do local da coleta de dados.

Caso você tenha alguma consideração ou dúvida sobre os aspectos éticos da pesquisa, poderá entrar em contato com o ***Comitê de Ética em Pesquisa da Universidade Guarulhos*** – Praça Tereza Cristina nº Centro Guarulhos.

Desde já agradecemos a sua colaboração.

Declaro que li e entendi os objetivos deste estudo, e que as dúvidas que tive foram esclarecidas pelo Pesquisador Responsável. Estou ciente que a participação é voluntária, e que, a qualquer momento tenho o direito de obter outros esclarecimentos sobre a pesquisa e de retirar a permissão para participar da mesma, sem qualquer penalidade ou prejuízo.

Nome do Sujeito de Pesquisa: _____

Assinatura do Sujeito de Pesquisa:_____

São Paulo, 28 de Abril de 2011.

__

_____

SONIA MARIA DA SILVA
Pesquisador Responsável

_____

Profa. Dra. Irai Cristina Boccato Alves
Orientador
Universidade de São Paulo-USP

ANEXO C: COMITE DE ETICA

UnG
CEPPE
Centro de Pós-Graduação e Pesquisa
Comitê de Ética em Pesquisa

Guarulhos, 23 de setembro de 2011.

Exmo(a). Sr(a).
Sonia Maria da Silva

Referência: **Aprovação de Projeto**
SISNEP/693 - "Estudo normativo da escala de avaliação de resultados de psicoterapia - OQ 45.2 na versão em português"

PARECER Nº 137/11

O Comitê de Ética em Pesquisa da Universidade Guarulhos analisou o Projeto de Pesquisa de sua autoria - "Estudo normativo da escala de avaliação de resultados de psicoterapia - OQ 45.2 na versão em português" - SISNEP/693, na reunião de 30.08.2011, e no uso das competências definidas na Res. CNS 196/96, considerou o projeto acima **aprovado.**

As orientações abaixo devem ser consideradas pelo Pesquisador Responsável durante a realização da pesquisa, visando que a mesma se desenvolva respeitando os padrões éticos:

- O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado e deve receber uma cópia do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado.
- O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade pelo CEP que o aprovou, aguardando seu parecer, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de regime oferecido a um dos grupos da pesquisa que requeiram ação imediata.
- Eventuais modificações ou emendas e eventos adversos ao protocolo, devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas.
- Esclarecemos a necessidade da apresentação de relatório final até **30.09.12.**

Jumara Silvia Van De Velde
Comitê de Ética em Pesquisa
Coordenadora