



Marlene Sofia Ferreira da Cunha

Propriedades psicométricas do Emotional and Psychological Outcome (EPO-1): Validação portuguesa numa amostra clínica

Universidade do Minho
Escola de Psicologia





Universidade do Minho

Escola de Psicologia

Marlene Sofia Ferreira da Cunha

**Propriedades psicométricas do
Emotional and Psychological
Outcome (EPO-1): Validação
portuguesa numa amostra clínica**

Dissertação de Mestrado
Mestrado em Psicologia
Clínica e Psicoterapia de
Adultos

Trabalho efetuado sob a orientação do **Prof. Dr. João Tiago Oliveira** (Universidade do Minho), do **Prof. Dr. Miguel Gonçalves** (Universidade do Minho) e do **Prof. Dr. Luís Janeiro** (Universidade do Algarve)

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição
CC BY

<https://creativecommons.org/licenses/by/4.0/>

Agradecimentos

Aos meus orientadores, por todo o apoio e suporte durante esta trajetória, e por todos os ensinamentos que me ajudaram a chegar até aqui.

À minha família, em particular a minha mãe e irmão, que sempre me apoiaram e ajudaram durante todo o meu percurso. Sem vocês não seria possível ter chegado até aqui, obrigada por toda a paciência, carinho e motivação que sempre me deram. Obrigada por todas as palavras de conforto nos momentos difíceis, e por celebrarem comigo cada pequena vitória.

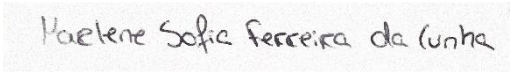
Aos meus amigos, que sempre me ouviram e confortaram quando precisei, e que me ajudaram todas as vezes sem hesitar. Um particular obrigada ao Duarte, por toda a ajuda e prontidão, e por ser um irmão de outra mãe.

Ao Tiago, por acreditar sempre em mim, por todas as palavras de carinho, pela (enorme) paciência e por ser o meu porto seguro. Tornaste esta jornada mais bonita!

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.



(Marlene Sofia Ferreira da Cunha)

Propriedades psicométricas do Emotional and Psychological Outcome (EPO-1): Validação portuguesa numa amostra clínica

Resumo

Os estudos empíricos demonstraram que a inclusão de medidas de *Routine Outcome Monitoring* (ROM) na prática clínica trazem benefícios para os clientes. Juntamente com a perceção vantajosa da sua inclusão no processo terapêutico, tem surgido também, ao longo do tempo, a necessidade de medidas ROM mais curtas e breves, de forma a facilitar o seu uso. O presente estudo tem como principal objetivo avaliar as propriedades psicométricas de uma medida de item único do bem-estar, o *Emotional and Psychological Outcome-1* (EPO-1). Para tal, foram analisadas 183 respostas a três instrumentos: o EPO-1, o *Outcome Questionnaire-10.2* (OQ-10.2) e *Outcome Rating Scale* (ORS) de uma amostra clínica portuguesa de 33 pacientes. Os resultados sugerem que o EPO-1 apresenta: uma boa validade discriminante, com um tamanho de efeito médio; uma boa validade convergente, com correlações elevadas e estatisticamente significativas tanto com a ORS ($r = .94$ na escala analógica e $r = .77$ na *Likert*, $p < .001$) bem como com o OQ-10.2 ($r = -.64$ na escala analógica e $r = -.61$ na *Likert*, $p < .001$); e, boa fidelidade teste-reteste, com um tamanho de efeito elevado ($r = .89$ na escala analógica e $r = .56$ na *Likert*, $p < .001$). Este estudo desempenhou um papel fundamental na validação de uma medida ROM de item único do bem-estar, estabelecendo uma base para investigações futuras. Deste modo, é importante o desenvolvimento de mais estudos em Portugal, dando ênfase à utilização de amostras mais representativas e diversas da população clínica e não clínica.

Palavras-chave: Bem-estar, Emotional and Psychological Outcome-1, Item Único, Propriedades Psicométricas, ROM

Psychometric properties of Emotional and Psychological Outcome (EPO-1): Portuguese validation on a clinical sample

Abstract

The empirical studies have demonstrated that the inclusion of Routine Outcome Monitoring (ROM) measurements in clinical practice bring benefits to the clients. Alongside the advantageous perception of its inclusion in the therapeutic process, there has also been, as time goes on, the need for shorter and briefer ROM measures, in order to facilitate its use. This study's main objective is to determine the psychometric properties of a single item measure of wellbeing, the Emotional and Psychological Outcome-1 (EPO-1). To that end, 183 responses to three instruments (the EPO-1, the Outcome Questionnaire-10.2 and the Outcome Rating Scale) were analyzed on a clinical sample of 33 Portuguese patients. The results suggest that the EPO-1 has: good discriminant validity, with an average effect size; good convergent validity, with high and statistically significant correlations on both ORS ($r = .94$ on the analogue scale and $r = .77$ on the Likert, $p < .001$) as well as OQ-10.2 ($r = -.64$ on the analogue scale and $r = -.61$ on the Likert, $p < .001$); and good test-retest fidelity, with high effect size ($r = .89$ on the analogue scale and $r = .56$ on the Likert, $p < .001$). This study was a key factor in the validation of a single item ROM measure of wellbeing, establishing a foundation for future research. With this in mind, it is important that more studies are conducted in Portugal, focusing on the usage of more diverse and representative samples of the clinical and non-clinical population.

Keywords: Wellbeing, Emotional and Psychological Outcome-1, Single Item, Psychometric Properties, ROM

Índice

Introdução	8
Método	12
Participantes	12
Instrumentos	12
Procedimentos.....	13
Resultados	15
Discussão	16
Referências	19

Índice de Tabelas

Tabela 1 <i>Teste t para comparação de médias entre momentos (inicial vs. final)</i>	15
--	----

Introdução

A psicoterapia é eficaz, tal como tem sido demonstrado por uma diversidade de estudos meta-analíticos (Barkham & Lambert, 2021). Estes estudos sugerem que o tamanho do efeito comparado com ausência de tratamento é de .75 a .85 (Wampold & Imel, 2015). Lutz et al. (2021) compararam estudos sobre psicoterapia com estudos sobre diversos tratamentos médicos e verificaram que a psicoterapia apresenta tamanhos de efeito maiores que, por exemplo, o tratamento com corticosteróide para asma crónica, com um tamanho do efeito de .56, ou a utilização de aspirina para prevenir doenças cardiovasculares, com um tamanho do efeito de .12.

Inicialmente, a eficácia da psicoterapia foi avaliada através de medições pré-pós tratamento, sendo desta forma possível avaliar se determinado tratamento é eficaz ou não. Todavia, existem outras formas de avaliar o sucesso terapêutico, tais como medições repetidas e até mesmo medições longitudinais intensivas. Este tipo de medições (de forma regular) é mais vantajosa, no sentido em que permite obter informações mais detalhadas sobre o tratamento em curso. Não seria possível recolher tais informações através de medições pré-pós tratamento, pois através das mesmas é apenas possível perceber se o tratamento funciona, podendo deixar muitas dúvidas sobre os casos de *dropout* e os casos de insucesso terapêutico, bem como qual é a evolução dos resultados do tratamento (Lutz et al., 2022).

Num artigo marcante publicado no *American Psychologist*, Howard et al. (1996) sugeriram a ideia de que uma monitorização constante do progresso do cliente ao longo do processo psicoterapêutico poderia melhorar o nosso conhecimento sobre a mudança em psicoterapia. Neste estudo, os autores introduziram um novo paradigma, a “investigação focada no paciente”, enfatizando a monitorização frequente dos resultados terapêuticos dos clientes utilizando medidas repetidas breves. Os autores desenvolveram também uma nova abordagem chamada “*patient profiling*”, que se baseia em duas teorias (o modelo psicoterapêutico de dosagem e o modelo psicoterapêutico de fases) e permite ao terapeuta analisar o progresso atual comparativamente com o progresso estimado de um paciente. Ou seja, o *patient profiling* permite verificar se os resultados psicoterapêuticos do cliente vão ao encontro da trajetória esperada, ajudando a identificar possíveis pacientes em risco.

A partir deste estudo começaram a surgir na literatura diversos artigos a abordar o tema, bem como revisões sistemáticas e meta-análises (Barkham & Broglio, 2023). Esta abordagem tem sido apresentada na literatura através de diferentes termos (tais como *measurement-based care*, *progress monitoring* e *feedback-informed treatment*), sendo um dos mais conhecidos *Routine Outcome*

Monitoring (ROM; Barkham et al., 2023). O ROM pode ser definido como o uso de medidas de autorrelato sobre o estado de saúde mental do paciente no começo do tratamento, e a sua repetição regular durante o processo de terapia (Lambert & Harmon, 2018). Barkham et al. (2023) sugerem que o ROM pode ser dividido em três fases: recolha de dados, *feedback*, e adaptação da terapia em função do *feedback*. A recolha de dados é realizada através de instrumentos objetivos de medição de resultados. Por sua vez, o *feedback* consiste em devolver ao cliente e/ou terapeuta informação sobre os dados recolhidos. O último passo consiste então em introduzir possíveis alterações necessárias na terapia tendo em conta os dados obtidos na recolha e o *feedback* discutido em sessão com o cliente.

Segundo Lambert e Harmon (2018), o ROM é suportado pela evidência empírica e é recomendado pela *American Psychological Association* (APA Presidential Task Force on Evidence-Based Practice, 2006). A utilização cuidadosa de um *feedback* sistemático pode facilitar a melhoria nos resultados terapêuticos dos clientes, especialmente na redução sintomática e no aumento do bem-estar (Bovendeerd et al., 2022). Nesse sentido, esta avaliação de resultados rotineira é importante pois permite aos terapeutas terem *feedback* sobre o estado atual dos clientes, se os mesmos estão ou não a progredir e, em caso afirmativo, de que forma o fazem, ou se não estão, porquê. Permite moldar o tratamento às necessidades do cliente e ao seu progresso terapêutico, ou falta dele, bem como avaliar se o tratamento está a ser eficaz (Lambert & Harmon, 2018). Ou seja, o ROM é útil para perceber os clientes que não estão a melhorar (Lambert, 2017). Os dados das medidas também são relevantes para o terapeuta devolver o *feedback* ao cliente e conversar com ele de forma colaborativa sobre o que está a acontecer em terapia (Lambert & Harmon, 2018).

Em 2003, Lambert et al. realizaram uma meta-análise com três estudos sobre o ROM. Os autores perceberam que o ROM permitia reduzir as taxas de deterioração e melhorava os resultados terapêuticos nos clientes com uma baixa resposta inicial ao tratamento, sugerindo que os terapeutas devem adotar formas de ROM na sua prática clínica. Num estudo mais recente de Barkham et al. (2023), foi possível confirmar que a utilização do ROM tem efeitos estatisticamente significativos nos resultados dos clientes. Esses efeitos, apesar de estatisticamente pequenos, são aditivos, ou seja, são efeitos que melhoram os resultados terapêuticos dos clientes para além da melhoria observada pelo próprio processo terapêutico em si. Corroborando o estudo anterior, Barkham e Broglia (2023) sugerem que o ROM pode melhorar os resultados terapêuticos em até 8%. Outra vantagem é que, independentemente do modelo teórico utilizado nas sessões de terapia, o ROM pode ser utilizado (Barkham & Broglia, 2023).

É possível, ainda, destacar uma outra vantagem do ROM, sendo esta o facto de que ele permite analisar de uma forma mais contínua e regular a eficácia dos terapeutas. Barkham e Broglia (2023), que reuniram questões e temas emergentes na literatura sobre a utilização do ROM, verificaram que uma das perguntas que surge por parte dos terapeutas é sobre a natureza da necessidade de uma avaliação dos resultados do cliente, em vez de ser realizada por parte de clínicos com anos de experiência em terapia, e que acreditam que são capazes de perceber se um cliente está a ter um bom progresso ou não. Contudo, a evidência sugere que os terapeutas têm tendência para se autoavaliarem de forma excessivamente positiva, mas nem sempre realista (Lambert, 2017; Pinner & Kivlighan, 2018). Com a utilização do ROM, torna-se possível para os terapeutas analisarem de forma mais objetiva a sua própria eficácia (Lambert, 2017; Pinner & Kivlighan, 2018).

Existem vários sistemas para monitorizar resultados terapêuticos (Boswell et al., 2015), tais como o *Treatment Outcome Package* (TOP; Kraus et al., 2005) ou o *Clinical Outcomes in Routine Evaluation (CORE) system* (Barkham et al., 2001, 2006). Entre os vários sistemas o *Outcome Questionnaire (OQ)-System* (Lambert et al., 1998) e o *Partners for Change Outcome Management System* (PCOMS; Duncan & Reese, 2015; Miller et al., 2005) são os mais utilizados. Esses sistemas são uma prática suplementar às intervenções já em curso e podem ser aplicados numa variedade de contextos (Brattland et al., 2018).

No estudo de Brattland et al. (2018) foi utilizado o *Partners for Change Outcome Management System* (PCOMS; Duncan & Reese, 2015; Miller et al., 2005), que consiste na *Outcome Rating Scale* (ORS; Miller et al., 2003) e na *Session Rating Scale* (SRS; Duncan et al., 2003), como forma de ROM. Este estudo demonstrou que a utilização do ROM foi benéfica para os clientes, que apresentaram maior probabilidade para uma melhoria sintomatológica e funcional, independentemente do seu grau inicial de sofrimento. No entanto, o ROM não possui efeitos positivos em todos os clientes, sendo necessário prestar atenção a aspetos da sua implementação como, por exemplo, a forma como é explicado ao cliente ou o ambiente em que está a ser utilizado. Neste sentido, o ROM deve ser um processo integrado na terapia, que faça sentido para o cliente e que lhe seja explicado de forma clara (Barkham et al., 2023). Segundo Lambert e Harmon (2018):

Ao contrário da implementação de um novo protocolo de tratamento onde as exigências são muitas e o treino é extensivo, a implementação de uma prática de ROM é comparativamente simples e, portanto, relativamente fácil para um clínico individual integrar a prática. (p. 2)

Para a implementação do ROM, é necessário que os terapeutas introduzam aos clientes a prática de relatarem regularmente o seu estado mental, compreendam o que está a ser avaliado, interpretem o *feedback* e abordem o progresso do cliente quando necessário. No entanto, alguns estudos alertam para possíveis dificuldades na implementação de um sistema de monitorização de resultados. Algumas características dos pacientes, como determinados sintomas, podem tornar-se uma barreira na implementação do ROM. O mesmo pode acontecer se os clínicos não tiverem confiança na medição dos dados, ou as organizações não possuírem capacidade tecnológica e financeira suficiente (Lewis et al., 2019). Também se podem encontrar outros obstáculos, tais como medo e desconfiança por parte dos terapeutas com a confidencialidade dos dados e com uma possível avaliação do seu desempenho, e privacidade e ética, uma vez que a recolha deste tipo de dados pode conter alguma informação que identifique os clientes, e por isso deve ser realizada de forma ética (Boswell et al., 2015). Uma das preocupações dos clínicos quando abordados sobre o ROM está relacionada com o “fardo da medição”, no sentido em que realizar medições de forma rotineira pode tornar-se um “fardo” para os clientes que têm constantemente de preencher questionários (Barkham & Broglia, 2023).

Para uma adoção eficaz do ROM, Barkham et al. (2023) sugerem a identificação de possíveis obstáculos antes da implementação, a atenção à sintomatologia e adequação de uma avaliação rotineira dos pacientes e a utilização de medidas clínicas práticas e apropriadas culturalmente. Por sua vez, Lewis et al. (2019) fornecem uma série de sugestões, entre as quais analisar momentos críticos para alcançar melhores resultados terapêuticos e utilizar medidas breves e psicometricamente robustas. Uma forma de reduzir o “fardo” das medições referido anteriormente, seria com a adoção de medidas de item único na avaliação de resultados. Essas medidas permitem fornecer uma só resposta ao conteúdo avaliado, abordando-o de forma unitária e direcionada ao item disposto. São medidas mais breves e parcimoniosas e, portanto, podem ser uma opção facilitadora para implementação do ROM, comportando menores custos financeiros, uma maior facilidade tecnológica e praticidade clínica (Patrician, 2004).

Segundo Loo (2002), os itens únicos são vantajosos quando se trata de reduzir o tempo de administração das medidas. Mais recentemente Allen et al. (2022) corroboram essa informação, mencionando o curto tempo de administração das medidas de item único. Abordam também outros aspetos positivos das mesmas, como a sua utilidade e eficiência, a facilidade de resposta por parte de quem as preenche e os baixos custos (associados ao processamento dos dados). Segundo Sarstedt e Wilczynski (2009), no que diz respeito aos aspetos práticos, tais como custos, taxa de resposta,

desenvolvimento da escala, fadiga mental e flexibilidade, os itens únicos são mais adequados que os itens múltiplos. No entanto, a utilização de medidas com múltiplos itens tem sido a norma devido às suas propriedades psicométricas (Loo, 2002). Itens únicos estão em desvantagem relativamente à confiabilidade (Allen et al., 2022; Loo, 2002; Sarstedt & Wilczynski, 2009), vulnerabilidade a erros de medida, e apreciação na investigação académica (Sarstedt & Wilczynski, 2009). No que diz respeito à validade preditiva, os itens únicos conseguem igualar o desempenho das escalas com vários itens apenas em situações muito particulares (Diamantopoulos et al., 2012). Nesse sentido, a utilização de medidas de item único deve ser realizada com cuidado (Diamantopoulos et al., 2012).

Tendo em vista as vantagens da implementação de uma monitorização rotineira de resultados e a necessidade do desenvolvimento de uma medida de avaliação de resultados prática e breve, foi criado um Consórcio Internacional - *the European Psychotherapy Consortium* (EPoC) – cujo um dos objetivos é o desenvolvimento de uma medida de item único sobre o bem-estar que reflita o progresso terapêutico do cliente ao longo das sessões e que possa ser utilizada em vários contextos e países. O EPoC é constituído por uma diversidade de investigadores que estão a realizar recolha de dados utilizando um item único em diferentes países, e, para alcançar este objetivo, serão seguidos três passos principais: introdução de um item único nos procedimentos de recolha de dados, desenvolver conexões entre diferentes medidas de resultado e recolha unificada de dados em *routine care*. O item único adotado pelo EPoC baseia-se no trabalho de Orlinsky e Howard (1986), denomina-se *Emotional and Psychological Outcome-1* (EPO-1) e até à presente data está traduzido em 11 línguas. O presente estudo enquadra-se no âmbito do EPoC e tem como principal objetivo avaliar as propriedades psicométricas do item único desenvolvido.

Método

Participantes

A amostra deste estudo é constituída por 183 respostas de 32 pacientes, com idades compreendidas entre 21-61 anos ($M = 26.68$ anos; $DP = 9.42$ anos) e de ambos os sexos. Cada paciente tem entre uma e 44 sessões, desde sessões de avaliação (sessão zero) até à sessão 44, e a sessão nove de média ($DP = 10.86$). A amostra é constituída tanto por participantes que já terminaram a terapia como por terapias ainda em curso. Os pacientes foram acompanhados por 10 terapeutas ou co-terapeutas.

Instrumentos

Neste estudo foram utilizados três instrumentos de recolha de dados: o *Emotional and Psychological Outcome-1* (EPO-1; Orlinsky & Howard, 1986), o *Outcome Questionnaire-10.2* (OQ-10.2; Lambert et al., 1998) e a *Outcome Rating Scale* (ORS; Miller et al., 2003).

Emotional and Psychological Outcome-1 (EPO-1)

O primeiro instrumento é a versão portuguesa baseada no item único de Orlinsky e Howard (1986), que consiste na seguinte pergunta “Neste momento, como é que se tem sentido em termos emocionais e psicológicos?” e possui opção de resposta em escala *Likert* (0 - muito mal: mal consigo lidar com as coisas; 1 - mal: a vida é muito difícil para mim por vezes; 2 - mais ou menos: com algum esforço consigo lidar com as coisas; 3 - bem: tenho os meus altos e baixos; e 4 - muito bem: não tenho queixas significativas) e numa escala analógica que vai de 0 (muito mal) até 100 (muito bem).

Outcome Questionnaire-10.2 (OQ-10.2; Lambert et al., 1998)

O OQ-10.2 (Lambert et al., 1998) é uma versão mais breve do OQ-45.2 (Lambert et al., 1996) com 10 itens, uma consistência interna de $\alpha=.87$ (Goates-Jones & Hill, 2008) e fidelidade teste-reteste de $r = .62$. Na população portuguesa os valores de consistência interna e fidelidade teste-reteste são $\alpha=.81$ e $r = .82$, respetivamente (Oliveira et al., 2020). Este questionário é composto por cinco itens que avaliam o bem-estar psicológico e cinco itens relacionados com o sofrimento psicológico. Quanto maior for a pontuação total, maior será o nível de sofrimento sintomatológico.

Outcome Rating Scale (ORS; Miller et al., 2003)

A ORS é uma escala breve com quatro itens em opção de resposta analógica, com uma consistência interna entre $\alpha=.87$ e $\alpha=.96$ e uma fidelidade teste-reteste de $r = .66$ numa segunda administração (Miller et al., 2003). Avalia os resultados do cliente no âmbito individual (bem-estar pessoal), interpessoal (família e relações chegadas), social (trabalho, escola e amigos) e global (sensação geral de bem-estar). Quanto maior for a pontuação total, melhores serão os resultados terapêuticos do cliente.

Procedimentos

Os dados dos participantes estão protegidos pela Comissão de Ética para a Investigação em Ciências Sociais e Humanas (CEICSH) da Universidade do Minho. A aprovação para a recolha de dados foi previamente realizada no projeto de investigação intitulado “*Eficácia Psicoterapêutica do Protocolo Unificado para o Tratamento Transdiagnóstico das Perturbações Emocionais em Contexto Comunitário*”

e foi concedida pela Subcomissão de Ética para as Ciências Sociais e Humanas da Universidade do Minho (SECSH 011/2018). A amostra foi recolhida entre dezembro de 2023 e abril de 2024 através da inclusão do EPO-1 em duas clínicas universitárias. Os participantes além de preencherem os questionários de sessão normalmente administrados pelas mesmas, preencheram também o item anteriormente referido no final de cada sessão. Os critérios de inclusão deste estudo são os mesmos utilizados pelas clínicas universitárias onde decorreu a recolha da amostra, e consistem em Indivíduos adultos que cumprem critérios para uma perturbação emocional, assinem o consentimento informado, preencham os questionários utilizados pelas clínicas e autorizem a gravação em vídeo das sessões.

Foram excluídos os participantes com diagnóstico de perturbação da personalidade, que tinham comorbilidade com outro tipo de perturbações (perturbações de consumo de substâncias, perturbações sexuais, perturbações alimentares, etc), apresentavam ideação suicida severa, sintomas psicóticos ou diagnóstico de perturbação bipolar.

Tradução

O procedimento de tradução do EPO-1 para a língua portuguesa consistiu nos seguintes passos principais: descrição geral da versão inglesa em termos simples; tradução do item para a língua portuguesa; retroversão do item de Português para Inglês e comparação das duas versões; desenvolvimento a partir desta comparação de quatro versões diferentes; avaliação da adequação das diversas opções (e.g., opinião de peritos e de pacientes); escolha da decisão mais clara do ponto de vista dos utilizadores; e apresentação da versão final do item ao Consórcio. Este processo foi conduzido por diferentes investigadores e profissionais de psicologia, de acordo com os procedimentos da *International Test Commission*.

Análise de dados

Após a recolha da amostra foi realizada a análise dos dados utilizando o *Software* IBM SPSS (versão 28.0). O primeiro passo foram os procedimentos de estatística descritiva dos dados, tendo sido possível observar que a mediana do número de sessões correspondia à sessão seis. Com base nesse resultado e para efeitos de estatística, todas as respostas que foram obtidas até à sessão seis foram interpretadas como uma sessão inicial, e após a mesma foram interpretadas como uma sessão final. De seguida avaliou-se a validade discriminante do EPO-1, por meio da análise das respostas ao mesmo numa sessão inicial comparativamente com uma sessão final, através do teste estatístico *t student* e com recurso ao *d de Cohen*. Foi avaliada também a validade convergente do EPO-1, através de

correlações com o OQ-10.2 e a ORS, com recurso ao *r de Pearson*. Segundo Cronbach e Meehl (1955), se os valores destas correlações forem elevados significa que existe boa validade convergente. O próximo passo foi a análise da fidelidade teste-reteste do EPO-1 (em ambas as opções de resposta) entre um momento inicial e um mais avançado, também com recurso ao *r de Pearson*. A última etapa foi a realização de correlações entre as duas escalas de resposta (*Likert* e analógica) do EPO-1 e entre as mesmas e o n° da sessão, ambas com recurso ao *r de Pearson*. Foram utilizados os critérios definidos por Cohen (1988), considerando-se o tamanho de efeito baixo se o valor estiver entre .20 e .49, moderado entre .50 e .79 e grande se for igual ou superior a .80. Para o *r de Pearson* foram utilizados os critérios definidos por Pearson (1895), sendo os valores entre .10 e .30 correspondentes a um efeito baixo, entre .30 e .50 efeito médio e igual ou superior a .50 um efeito grande.

Resultados

Validade discriminante

Nas sessões iniciais os participantes apresentam uma pontuação média inferior no EPO-1 ($M = 53.86$, $DP = 18.96$ na escala analógica; $M = 2.00$, $DP = .81$ na escala *Likert*) comparativamente com as sessões finais ($M = 63.03$, $DP = 18.73$ na escala analógica; $M = 2.57$, $DP = .84$ na escala *Likert*). As diferenças entre estes valores são estatisticamente significativas ($p < .001$) e apresentam um tamanho de efeito médio (ver Tabela 1).

Tabela 1

Teste t para comparação de médias entre momentos (inicial vs. final)

	Inicial		Final		<i>t(gf)</i>	<i>p</i>	<i>d</i> <i>Cohen</i>
	<i>M</i>	<i>DP</i>	<i>M</i>	<i>DP</i>			
EPO-1 (0-100)	53.86	18.96	63.03	18.73	-3.157(169)	<.001	.49
EPO-1 Likert (0-4)	2.00	.81	2.57	.84	-4.312(158)	<.001	.69

Nota. M = Média; DP = Desvio Padrão; *t* = Teste *t student*

Validade convergente

O EPO-1 está positivamente correlacionado com a ORS, tanto na escala analógica ($r = .94$, $p < .001$) como na escala de *Likert* ($r = .77$, $p < .001$). Ambas as correlações são estatisticamente significativas e elevadas. Relativamente ao OQ-10.2, o EPO-1 está negativamente correlacionado com o mesmo, tanto na escala analógica ($r = -.64$, $p < .001$) como na escala *Likert* ($r = -.61$, $p < .001$), com ambas as correlações estatisticamente significativas e elevadas. Embora as correlações apresentem sentidos opostos, indicam que quanto maiores as pontuações no EPO-1, melhores os resultados. Além disso, a magnitude do efeito é maior quando se utiliza a escala *Likert*.

Fidelidade teste-reteste

O EPO-1 apresenta boa fidelidade teste-reteste (feita em 28 casos). Verificou-se um tamanho de efeito elevado tanto na escala analógica ($r = .89$, $p < .001$) como na escala *Likert* ($r = .56$, $p < .001$) e ambos os valores são estatisticamente significativos.

Correlações

As escalas de resposta do EPO-1 estão positivamente correlacionadas entre si ($r = .62$, $p < .001$), de forma estatisticamente significativa e com um tamanho de efeito elevado. Estão também correlacionadas de forma positiva com o n° da sessão ($r = .31$, $p < .001$ na escala analógica; $r = .28$, $p < .001$ na escala *Likert*). Estes valores indicam uma baixa correlação, no entanto estatisticamente significativa em ambas.

Discussão

Este estudo está incluído no EPoC, e o seu objetivo foi o de descrever as propriedades psicométricas de um item único desenvolvido no domínio do mesmo. Os principais resultados deste estudo demonstram que o bem-estar dos clientes é maior conforme o n° de sessões aumenta, que o EPO-1 está correlacionado tanto com o OQ-10.2 como com a ORS, e as opções de resposta (analógica e *Likert*) estão correlacionadas entre si. Deste modo, estes resultados sugerem que o EPO-1 possui boas qualidades psicométricas.

Tendo em vista as vantagens do ROM e a necessidade do desenvolvimento de uma medida de avaliação de resultados prática e breve, que pudesse ser partilhada pelos diferentes países envolvidos no EPoC, foi desenvolvido um item único – o EPO-1 (Orlinsky & Howard, 1986) – cujos resultados apontam para boas propriedades psicométricas, nomeadamente boa validade discriminante e convergente, boa fidelidade teste-reteste, bem como sensibilidade à mudança. Estes resultados, apesar de exploratórios, vão no sentido de validar a ideia de que os itens únicos não estão em desvantagem

comparativamente com os itens múltiplos em termos de propriedades psicométricas, como referido por Allen et al. (2022), Loo (2002) e Sarstedt e Wilczynski (2009). É possível observar que as pontuações no EPO-1 foram estatisticamente superiores numa sessão final comparativamente com uma sessão inicial e que o mesmo está correlacionado com o nº da sessão (apesar de ser uma correlação baixa), o que indica que o bem-estar dos clientes é maior conforme aumenta o nº da sessão. Estes resultados sugerem que o EPO-1 é sensível à mudança sintomatológica ao longo do tempo.

O EPO-1 está positivamente correlacionado com a ORS. A ORS é um instrumento que mede os resultados individuais, interpessoais, sociais e globais do cliente, que envolvem construtos como bem-estar pessoal e sensação global de bem-estar. Relativamente ao OQ-10.2, uma medida breve que avalia o bem-estar/sofrimento psicológico, e cujas pontuações elevadas significam maiores níveis de sofrimento psicológico, está negativamente correlacionado com o EPO-1. Como é possível observar nos resultados, ambos os valores destas correlações são elevados, o que demonstra uma boa validade convergente por parte do EPO-1. A partir dos resultados é possível também verificar que o EPO-1 apresenta bons valores de fidelidade teste-reteste. Comparando com outras medidas breves de avaliação, como a ORS e o OQ-10.2, que apresentam valores de fidelidade teste-reteste de $r = .66$ e $r = .82$, respetivamente, é possível observar que a escala analógica do EPO-1 possui um valor superior enquanto a escala *Likert* apresenta um valor inferior. Relativamente a outros instrumentos, como por exemplo o *Brief Symptom Inventory 18* (BSI 18; Derogatis, 2001), uma medida breve de 18 itens que avalia a sintomatologia psicológica em populações comunitárias e clínicas, apresenta valores moderados a altos na população clínica portuguesa, entre $r = .36$ e $r = .58$ (Nazaré et al., 2017).

Este estudo contribuiu para a validação de um item único sobre o bem-estar que reflita o progresso terapêutico do cliente ao longo das sessões em Portugal, o que apresenta vantagens para a área clínica: baixos custos financeiros, facilidade de resposta por parte dos clientes, facilidade tecnológica, praticidade clínica, curto tempo de administração e baixo nível de fadiga mental por parte dos clientes. Contribuiu ainda para a facilitação de uma recolha de dados unificada em *routine care*; bem como a possibilidade de recolher dados em diferentes países utilizando uma só medida (uma vez que este item único está a ser também traduzido e validado em diferentes países), o que permite comparar resultados entre diferentes locais do mundo.

Limitações

Este estudo apresenta algumas limitações, nomeadamente a natureza hierárquica da amostra e a recolha de dados realizada apenas em duas clínicas universitárias, o que não é suficientemente

representativo da população clínica em Portugal. Além disso, a amostra não abrange toda a diversidade da população clínica, uma vez que foram excluídos indivíduos com várias perturbações mentais e as suas diferentes sintomatologias. O desenho de investigação do presente estudo não permitiu que fossem realizadas análises mais aprofundadas, e desse modo não foi possível verificar se a melhoria no bem-estar dos clientes foi potencializada pelos efeitos do ROM ou se essa melhoria se deve apenas ao processo terapêutico em si. Para estudos futuros recomenda-se a recolha de amostra não clínica, para possíveis correlações entre a mesma e a amostra clínica.

Conclusão

Este estudo validou o EPO-1 como uma medida psicométrica eficaz para avaliar o bem-estar dos clientes ao longo das sessões terapêuticas. Apesar das limitações da amostra, o EPO-1 demonstrou ser uma ferramenta prática e eficiente para monitorar o progresso terapêutico, podendo contribuir para a melhoria dos resultados clínicos e facilitando a recolha de dados em diversos contextos. Deste modo é possível concluir que a presente investigação contribuiu para os objetivos do EPoC, particularmente para a expansão do conhecimento científico existente em psicoterapia.

Referências

- Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science: A call to action. *European Journal of Psychological Assessment*, 38(1), 1-5. <https://doi.org/10.1027/1015-5759/a000699>
- Barkham, M., & Broglia, E., on behalf of The SCORE Consortium (2023). Routine outcome monitoring (ROM) and feedback in university student counselling and mental health services: Considerations for practitioners and service leads. *Counselling and Psychotherapy Research*, 00, 1-13. <https://doi.org/10.1002/capr.12694>
- Barkham, M., Jong, K. D., Delgadillo, J., & Lutz, W. (2023). Routine outcome monitoring (ROM) and feedback: Research review and recommendations. *Psychotherapy Research*, 33(7), 841-855. <https://doi.org/10.1080/10503307.2023.2181114>
- Barkham, M., & Lambert, M. J. (2021). The efficacy and effectiveness of psychological therapies. In M. Barkham, W. Lutz & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change: 50th anniversary edition* (7th ed., pp. 135-189). John Wiley & Sons, Inc.
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service profiling and outcomes benchmarking using the core-om: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69(2), 184–196. <https://doi.org/10.1037/0022-006X.69.2.184>
- Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the core-om and core system. *Counselling & Psychotherapy Research*, 6(1), 3-15. <https://doi.org/10.1080/14733140600581218>
- Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, 25(1), 6-19. <https://doi.org/10.1080/10503307.2013.817696>
- Bovendeerd, B., Jong, K. D., Groot, E. D., Moerbeek, M., & Keijser, J. D. (2022). Enhancing the effect of psychotherapy through systematic client feedback in outpatient mental healthcare: A cluster randomized trial. *Psychotherapy Research*, 32(6), 710-722. <https://doi.org/10.1080/10503307.2021.2015637>

- Brattland, H., Koksvik, J. M., Burkeland, O., Gråwe, R. W., Klöckner, C., Linaker, O. M., Ryum, T., Wampold, B., Lara-Cabrera, M. L., & Iversen, V. C. (2018). The effects of routine outcome monitoring (rom) on therapy outcomes in the course of na implementation process: A randomized clinical trial. *Journal of Counseling Psychology*, 65(5), 641–652. <https://doi.org/10.1037/cou0000286>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>
- Derogatis, L. R. (2001). *BSI 18 – Brief Symptom Inventory 18: Administration, scoring, and procedures manual*. Minneapolis, MN: Pearson.
- Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, 40, 434–449. <https://doi.org/10.1007/s11747-011-0300-3>
- Duncan, B. L., Miller, S. D., Sparks, J. A., Claud, D. A., Reynolds, L. R., Brown, J., & Johnson, L. D. (2003). The Session Rating Scale: Preliminary psychometric properties of a “working” alliance measure. *Journal of Brief Therapy*, 3(1), 3–12.
- Duncan, B. L., & Reese, R. J. (2015). The Partners for Change Outcome Management System (pcoms) revisiting the client’s frame of reference. *Psychotherapy*, 52(4), 391–401. <https://doi.org/10.1037/pst0000026>
- Goates-Jones, M., & Hill, C. E. (2008). Treatment preference, treatment-preference match, and psychotherapist credibility: Influence on session outcome and preference shift. *Psychotherapy: Theory, Research, Practice, Training*, 45(1), 61–74. <https://doi.org/10.1037/0033-3204.45.1.61>
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, 51(10), 1059–1064. <https://doi.org/10.1037/0003-066X.51.10.1059>

- Kraus, D. R., Seligman, D. A., & Jordan, J. R. (2005). Validation of a behavioral health treatment outcome and assessment tool designed for naturalistic settings: The treatment outcome package. *Journal of Clinical Psychology, 61*(3), 285-314. <https://doi.org/10.1002/jclp.20084>
- Lambert, M. J. (2017). Maximizing psychotherapy outcome beyond evidence-based medicine. *Psychotherapy and Psychosomatics, 86*(2), 80-89. <https://doi.org/10.1159/000455170>
- Lambert, M. J., Finch, A. M., Okiishi, J., Burlingame, G. M., McKelvey, C., & Reisinger, C. W. (1998). *Administration and scoring manual for the OQ-10.2: An adult outcome questionnaire for screening individuals and population outcome monitoring*. American Professional Credentialing Services LLC.
- Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G., & Reisinger, C. W. (1996). *Administration and scoring manual for the Outcome Questionnaire (OQ-45.2)*. Wilmington, DE: American Professional Credentialing Services, 35.
- Lambert, M. J., & Harmon, K. L. (2018). The merits of implementing routine outcome monitoring in clinical practice. *Clinical Psychology: Science and Practice, 25*(4), Article e12268. <https://doi.org/10.1111/cpsp.12268>
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice, 10*(3), 288–301. <https://doi.org/10.1093/clipsy.bpg025>
- Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2019). Implementing measurement-based care in behavioral health: A review. *JAMA Psychiatry, 76*(3), 324-335. <https://doi.org/10.1001/jamapsychiatry.2018.3329>
- Loo, R. (2002). A caveat on using single-item versus multiple-item scales. *Journal of Managerial Psychology, 17*(1), 68-75. <https://doi.org/10.1108/02683940210415933>
- Lutz, W., Jong, K. D., Rubel, J. A., & Delgadillo, J. (2021). Measuring, predicting, and tracking change in psychotherapy. In M. Barkham, W. Lutz & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change: 50th anniversary edition* (7th ed., pp. 89-133). John Wiley & Sons, Inc.

- Lutz, W., Schwartz, B., & Delgadillo, J. (2022). Measurement-based and data-informed psychological therapy. *Annual Review of Clinical Psychology*, 18(1), 71-98. <https://doi.org/10.1146/annurev-clinpsy-071720-014821>
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The Outcome Rating Scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, 2(2), 91–100.
- Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology*, 61(2), 199 –208. <https://doi.org/10.1002/jclp.20111>
- Nazaré, B., Pereira, M., & Canavarro, M. C. (2017). Avaliação breve da psicossintomatologia: Análise factorial confirmatória da versão portuguesa do brief symptom inventory 18 (bsi 18). *Análise Psicológica*, 35(2), 213-230. <https://doi.org/10.14417/ap.1287>
- Oliveira, J. T., Ribeiro, A. P., & Gonçalves, M. M. (2020). Ambivalence in psychotherapy questionnaire: Development and validation studies. *Clinical Psychology & Psychotherapy*, 27(5), 727–735. <https://doi.org/10.1002/cpp.2457>
- Orlinsky, D. E., & Howard, K. I. (1986). The psychological interior of psychotherapy: Explorations with the Therapy Session Reports. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 477–501). Guilford Press.
- Patrician, P. A. (2004). Single-item graphic representational scales. *Nursing Research*, 53(5), 347-352. <https://doi.org/10.1097/00006199-200409000-00011>
- Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proceedings of the royal society of london*, 58, 240–242. <https://doi.org/10.1098/rspl.1895.0041>
- Pinner, D. H., & Kivlighan, D. M., III. (2018). The ethical implications and utility of routine outcome monitoring in determining boundaries of competence in practice. *Professional Psychology: Research and Practice*, 49(4), 247–254. <https://doi.org/10.1037/pro0000203>
- Sarstedt, M., & Wilczynski, P. (2009). More for less? A comparison of single-item and multi-item measures. *Die Betriebswirtschaft*, 69(2), 211-227.
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203582015>