



Ispá

Instituto Universitário
de Ciências Psicológicas,
Sociais e da Vida

**PRÁTICA DELIBERADA E DESENVOLVIMENTO DE
COMPETÊNCIAS INTERPESSOAIS FACILITADORAS: UM
ESTUDO LONGITUDINAL EM PSICOTERAPEUTAS EM
FORMAÇÃO**

ANDRÉ MENDONÇA DA COSTA

Orientador de Dissertação:

PROFESSOR DOUTOR DANIEL MONTEIRO CUNHA DE SOUSA

Professor de Seminário de Dissertação:

PROFESSOR DOUTOR DANIEL MONTEIRO CUNHA DE SOUSA

Dissertação submetida como requisito parcial para a obtenção do grau de:
MESTRE EM PSICOLOGIA
Especialidade em Psicologia Clínica

Dissertação de
Mestrado realizada sob a
orientação de Professor Doutor
Daniel Cunha Monteiro de
Sousa, apresentada no ISPA –
Instituto Universitário para
obtenção de grau de Mestre na
especialidade de Psicologia
Clínica

Agradecimentos

Primeiramente, expresso a minha profunda gratidão ao Professor Daniel Sousa pela orientação, pela disponibilidade constante e pela rapidez com que respondeu às minhas necessidades ao longo deste percurso.

À Margarida e à Paula, agradeço profundamente a dedicação, o apoio e o muito que aprendi ao longo do caminho que fizemos juntos neste trabalho.

Aos meus pais, à minha família e aos meus amigos, agradeço o amor, o apoio e a força que sempre me deram.

A todo o corpo docente do ISPA, deixo o meu sincero agradecimento. Cada unidade curricular, cada discussão em aula e cada contributo académico ajudaram a moldar a minha forma de pensar, a expandir horizontes e a desenvolver a minha capacidade crítica.

Aos alumni do ISPA, agradeço igualmente por todo o impacto que tiveram no meu crescimento e pela inspiração que trouxeram a este percurso.

Por último, um agradecimento muito especial à Turma 8, pela camaradagem e espírito de entreajuda que nasceu em plena pandemia e que se manteve até aos dias de hoje.

A todos, o meu mais sentido obrigado.

RESUMO

Problema: Embora a prática deliberada (PD) apresente eficácia reconhecida noutros domínios, o seu contributo específico para o desenvolvimento de competências clínicas em psicoterapeutas em formação continua pouco clarificado.

Objetivos: Avaliar se a PD promove melhorias nas competências interpessoais facilitadoras, analisar a congruência entre autoavaliação (FIS-SR) e heteroavaliação (FIS) e, de forma exploratória, examinar a evolução da discrepância entre ambas ao longo do tempo.

Método: Conduziu-se um estudo quantitativo de natureza pré-experimental, utilizando as medidas Facilitative Interpersonal Skills Performance Task e Facilitative Interpersonal Skills Self-Report. As avaliações foram recolhidas em três momentos distintos da formação, permitindo comparar a progressão objetiva com a perceção subjetiva de competência.

Resultados: Identificou-se uma melhoria estatisticamente significativa das FIS ao longo dos três momentos, corroborando a eficácia da PD no reforço destas competências. As autoavaliações mantiveram-se consistentemente superiores às heteroavaliações, refletindo um padrão de sobrestimação. Observou-se ainda uma tendência exploratória de redução da discrepância entre medidas, sobretudo explicada pela melhoria objetiva captada pela FIS.

Conclusões: Os resultados sugerem que a prática deliberada constitui um método promissor na promoção das competências interpessoais facilitadoras, embora persistam desafios na calibração metacognitiva dos formandos. Investigações futuras deverão recorrer a amostras mais amplas e delineamentos mais controlados, explorando ainda variáveis adicionais que possam influenciar o desenvolvimento destas competências.

Palavras-Chave: Prática Deliberada; Competências Interpessoais Facilitadoras; Formação em Psicoterapia. Autoavaliação de Competências; Avaliação Externa.

ABSTRACT

Problem: Although deliberate practice (DP) has demonstrated effectiveness in other domains, its specific contribution to the development of clinical skills in trainee psychotherapists remains insufficiently clarified.

Objectives: To assess whether DP promotes improvements in facilitative interpersonal skills, to examine the congruence between self-assessment (FIS-SR) and external assessment (FIS), and, in an exploratory manner, to evaluate how the discrepancy between the two evolves over time.

Method: A quantitative pre-experimental study was conducted using the Facilitative Interpersonal Skills Performance Task and the Facilitative Interpersonal Skills Self-Report. Assessments were collected at three distinct moments during training, allowing for the comparison of objective progression with the subjective perception of competence.

Results: A statistically significant improvement in FIS scores was observed across the three assessment moments, supporting the efficacy of DP in strengthening these competencies. Self-assessments remained consistently higher than external assessments, reflecting a pattern of overestimation. An exploratory trend toward a reduction in the discrepancy between measures was also noted, largely explained by the objective improvement captured by the FIS.

Conclusions: The findings suggest that deliberate practice is a promising method for enhancing facilitative interpersonal skills, although challenges persist regarding trainees' metacognitive calibration. Future research should employ larger samples and more controlled designs, while also exploring additional variables that may influence the development of these competencies.

Keywords: Deliberate Practice; Facilitative Interpersonal Skills; Psychotherapy Training; Competence Self-Assessment; External Evaluation.

Índice

Introdução	7
Método	4
Visão Geral	4
Participantes	5
Curso de Especialização em Psicoterapia	6
Instrumentos	6
Consentimento Informado e Questionário Sociodemográfico	6
Vídeos de Evocação Clínica	7
Facility Interpersonal Skills Performance Task	8
Facilitative Interpersonal Skills – Self Report (FIS-SR)	9
Procedimento	9
Análise de Dados	10
Resultados	12
Verificação do Pressuposto de Normalidade (FIS e FIS-SR)	12
FIS Global (Heteroavaliação)	12
Comparações da FIS (heteroavaliação) Entre Momentos (Itens 1 a 8)	14
Evolução da FIS-SR (Autoavaliação) ao Longo dos Três Momentos	15
Comparação da FIS (heteroavaliação) e FIS-SR (autoavaliação) ao Longo dos Três Momentos Avaliativos	17
Discussão	21
Limitações	23
Implicações Para Investigação Futura	24
Conclusão	25
Referências	26
Anexos	33

Revisão da Literatura	33
Consentimento Informado	62
Questionário Sociodemográfico.....	63
Facilitative Interpersonal Skills Rating Scale	65
Facilitative Interpersonal Skills – Self Report.....	71
Coeficiente de Correlação Intraclasse (CCI) (Output do SPSS).....	72
Teste de Normalidade dos Resíduos – FIS (Output do SPSS).....	73
Teste de Normalidade dos Resíduos – FIS-SR (Output do SPSS)	74
ANOVA Medidas Repetidas – FIS Global (Outputs do SPSS).....	75
Anova Medidas Repetidas – FIS – Itens 1 e 4 (Outputs do SPSS).....	75
Friedman – FIS – Itens 2, 3 e 5 a 8 (Outputs do SPSS).....	78
Wilcoxon – FIS – Itens 2, 3 e 5 a 8 (Outputs do SPSS)	80
ANOVA Medidas Repetidas – FIS-SR Global (Outputs do SPSS)	82
ANOVA Medidas Repetidas – FIS-SR – Itens 2, 3 e 8 (Outputs do SPSS).....	83
Friedman – FIS-SR – Itens 1, 4, 5, 6, e 7 (Outputs do SPSS)	86
Wilcoxon – FIS-SR – Itens 1, 4, 5, 6, e 7 (Outputs do SPSS).....	88
ANOVA Medidas Repetidas – 2 Fatores – FIS e FIS-SR Globais (Outputs do SPSS)	90
Wilcoxon – Comparação FIS e FIS-SR – Itens 1 a 8 (Outputs do SPSS).....	92

Abreviaturas

CCI = Coeficiente de Correlação Intraclasse

FIS = Facilitative Interpersonal Skills

PD = Prática Deliberada

Lista de Tabelas

Tabela 1. Dados sociodemográficos dos participantes (N = 17)

Tabela 2. Estatísticas Descritivas e Resultados Principais da ANOVA de Medidas Repetidas Para a Pontuação Global da FIS

Tabela 3. Comparações pairwise (Bonferroni) para as médias da FIS Global nos três momentos de avaliação

Tabela 4. Resultados do teste de Wilcoxon para a comparação entre heteroavaliação (FIS) e autoavaliação (FIS-SR) nos três momentos

Lista de Figuras

Figura 1. Evolução das médias da FIS global ao longo dos três momentos de avaliação

Figura 2. Comparação das médias das competências FIS ao longo dos três momentos de avaliação

Figura 3. Médias da pontuação global da FIS-SR nos três momentos de avaliação

Figura 4. Evolução das médias globais da FIS (heteroavaliação) e da FIS-SR (autoavaliação) nos três momentos avaliativos

Introdução

A evidência acumulada ao longo de várias décadas indica de forma consistente que a psicoterapia é eficaz para múltiplas perturbações psicológicas, com melhorias substancialmente superiores às verificadas em grupos não tratados (Lambert, 2013; Smith & Glass, 1977; Wampold & Imel, 2015). Estes efeitos robustos surgem transversalmente a diferentes orientações teóricas, sugerindo que os fatores comuns desempenham um papel

central (Bohart & Tallman, 2010; Frank & Frank, 1991). Ademais, a variação entre terapeutas tem revelado que uma parte significativa dos resultados clínicos é atribuível a características individuais dos profissionais (Baldwin & Imel, 2013; Laska et al., 2013), justificando uma investigação aprofundada dos processos que sustentam a expertise terapêutica.

Estudos sugerem que variáveis estruturais como idade, gênero, modelo teórico ou anos de experiência predizem pouco os resultados terapêuticos (Goldberg et al., 2016; Wampold, 2015). Em contraste, terapeutas mais eficazes distinguem-se pelas suas competências relacionais e responsividade interpessoal (Heinonen & Nissen-Lie, 2020; Rogers, 1957; Truax & Carkhuff, 1967). Estas capacidades, frequentemente conceptualizadas como *Competências Interpessoais Facilitadoras* (FIS) incluem empatia, expressividade emocional, reparação de rupturas e promoção de expectativas positivas (Anderson et al., 2009; Norcross & Lambert, 2018). As FIS sustentam a formação de uma aliança terapêutica sólida, um dos preditores mais consistentes de eficácia (Horvath et al., 2011; Wampold & Imel, 2015).

Em termos operacionais, as FIS têm sido avaliadas sobretudo através da FIS Performance Task, que recorre a vídeos padronizados para simular situações terapêuticas exigentes. Este procedimento permite observar dimensões como fluência verbal, promoção de esperança, persuasão terapêutica, expressão emocional, aceitação, empatia, manutenção e reparação de aliança (Anderson et al., 2019), constituindo um marcador robusto da responsividade interpessoal (Rousmaniere, 2017). A investigação tem apontado de forma consistente que níveis mais elevados de FIS predizem melhores resultados clínicos (Anderson et al., 2009, 2016; Allen et al., 2023).

Estudos que comparam autoavaliação e desempenho observado evidenciam que os terapeutas tendem a sobrestimar as suas competências relacionais. Longley et al. (2023) observaram que as autoavaliações das FIS são sistematicamente superiores às classificações de avaliadores independentes, ainda que moderadamente correlacionadas. Este padrão converge com a literatura mais ampla sobre autoavaliação clínica, que identifica dificuldades metacognitivas persistentes e limitações da reflexão espontânea na produção de julgamentos sobre o próprio desempenho (Bennett-Levy, 2019; Macdonald & Mellor-Clark, 2015).

A investigação contemporânea sobre expertise tem aprofundado a compreensão destes processos, evidenciando que terapeutas altamente eficazes apresentam maior precisão,

responsividade e capacidade de automonitorização (Ericsson & Pool, 2016; Lambert & Shimokawa, 2011; Vaz et al., 2023). A literatura sobre desenvolvimento de competência profissional em psicólogos e psicoterapeutas sugere que o desempenho excepcional depende sobretudo da qualidade e estrutura do treino, incluindo prática deliberada, feedback especializado e critérios claros de desempenho, fatores que têm um impacto muito superior ao mero acumular de anos de experiência (Ericsson & Pool, 2016; Tracey et al., 2014). Esta evidência converge com trabalhos que destacam a importância da autorreflexão guiada e de oportunidades sistemáticas de ajustamento das competências clínicas (Bennett-Levy, 2019). Assim, compreender como os psicólogos em formação avaliam as suas competências torna-se fundamental, ainda que exista uma tendência inicial para sobrestimação, as discrepâncias entre autoavaliação e desempenho observado tendem a diminuir quando a formação inclui critérios explícitos, feedback contínuo e práticas sistemáticas de ajustamento (Longley et al., 2023; Macdonald & Mellor-Clark, 2015). Estudos qualitativos reforçam esta perspetiva, identificando barreiras cognitivas e estruturais à autoavaliação rigorosa e destacando a importância de modelos internos claros de competência (Banner et al., 2025; Bennett-Levy, 2019).

A prática deliberada (PD) destaca-se como método eficaz para o desenvolvimento de competências em diferentes contextos profissionais, centrando-se na repetição intencional de tarefas, no treino de componentes específicos e no feedback contínuo (Ericsson & Pool, 2016). Evidência recente, incluindo ensaios controlados, intervenções piloto e revisões sistemáticas, indicam que programas baseados em PD produzem ganhos mensuráveis em competências relacionais e processuais centrais (Clements-Hickman & Harris, 2024; Lagerberg et al., 2024; Larsson et al., 2023; Mahon, 2024; Nurse et al., 2024; Olsson et al., 2024). Alguns estudos apontam para benefícios consistentes, especialmente quando os programas incluem treino intensivo, tarefas específicas e feedback especializado (Brand et al., 2025; Nissen-Lie, 2025).

Apesar do interesse crescente, a eficácia da prática deliberada na psicoterapia permanece parcialmente inconclusiva. A evidência disponível, embora promissora, é ainda limitada e apresenta resultados heterogêneos, sugerindo ganhos modestos e nem sempre consistentes (Chow et al., 2015; Diamond et al., 2025; Goldberg et al., 2016). Neste sentido, Diamond et al. (2025) sublinham que, apesar de existirem sinais encorajadores, os resultados devem ser interpretados com cautela devido à variabilidade metodológica e à definição inconsistente da PD entre estudos.

A identificação das oito competências FIS como preditores robustos da eficácia clínica (Allen et al., 2023; Anderson et al., 2009; Anderson & Patterson, 2013), aliada à evidência emergente sobre modelos contemporâneos de PD, reforça a importância de integrar estas dimensões na formação de terapeutas. Assim, coloca-se a possibilidade de que programas estruturados segundo os princípios da PD contribuam para melhorias significativas no desempenho clínico (Allen et al., 2023; Brand et al., 2025; Larsson et al., 2023; Olsson et al., 2024).

Neste enquadramento, colocam-se duas hipóteses principais e uma exploratória: (H1) a prática deliberada conduz a uma melhoria estatisticamente significativa das competências interpessoais facilitadoras, avaliadas através da FIS (heteroavaliação), ao longo dos três momentos da formação; (H2) a autoavaliação das competências interpessoais facilitadoras (FIS-SR) mantém-se sistematicamente superior à heteroavaliação (FIS) em todos os momentos avaliativos; (H3 — exploratória) poderá observar-se uma tendência de diminuição da discrepância entre autoavaliação e heteroavaliação ao longo do tempo de formação.

Método

Visão Geral

O presente estudo adota uma metodologia de investigação de natureza quantitativa, adequada aos objetivos de testar as hipóteses previamente formuladas. Esta abordagem fundamenta-se na análise estatística de dados numéricos, permitindo gerar indicadores objetivos e representações gráficas que facilitam a interpretação dos resultados (Creswell & Creswell, 2018; Marôco, 2021).

O delineamento assume um carácter pré-experimental, na medida em que não existe grupo de controlo e os participantes estão expostos a outras atividades formativas ao longo do percurso, o que limita o controlo sobre possíveis fatores externos. Simultaneamente, assume um carácter longitudinal de medidas repetidas, uma vez que a evolução dos participantes foi acompanhada ao longo de três momentos de recolha: fevereiro de 2024 (pré-formação), maio de 2024 (pós-primeiro módulo de PD) e junho de 2025 (pós-segundo módulo de PD).

A recolha foi realizada na plataforma Skillsetter, onde os formandos submetiam as suas respostas. Em paralelo, cada participante enviou igualmente a sua autoavaliação por e-mail.

As respostas foram cotadas por três avaliadores independentes, para determinar a sua fiabilidade e grau de concordância, procedeu-se ao cálculo do coeficiente de correlação intraclasse (CCI).

Participantes

A Tabela 1 apresenta as estatísticas descritivas da amostra final, composta por 17 participantes (N = 17). O estudo teve início com 23 participantes, contudo, verificou-se a desistência de 1 participante durante o primeiro ano de formação e de 5 adicionais no segundo ano, totalizando seis *dropouts*. A amostra é constituída exclusivamente por psicólogos clínicos a frequentar um curso de especialização de psicoterapia. A maioria identificou-se como sendo do sexo feminino (82,4%), enquanto 17,6% identificaram-se como sendo do sexo masculino. A idade dos participantes variou entre 25 e 60 anos (M = 37,06; DP = 11,05). A experiência profissional situou-se entre 1 e 29 anos (M = 6,27; DP = 8,47), e o número de horas semanais de prática clínica variou entre 1 e 38 horas (M = 16,79; DP = 11,83).

Relativamente ao modelo teórico seguido na prática clínica, a maioria dos participantes reportou adotar um enquadramento integrativo (82,4%), enquanto os restantes se distribuíram pelos modelos existencial, psicodinâmico e cognitivo-comportamental (5,9% cada).

Tabela 1

Dados Sociodemográficos dos Participantes (N = 17)

	Categoria	N	%	Mín.	Máx.	Média	DP
Sexo	Feminino	14	82,4%	—	—	—	—
	Masculino	3	17,6%	—	—	—	—
Idade (anos)	—	—	—	25	60	37,06	11,05
Anos de experiência	—	—	—	1,0	29,0	6,27	8,47
Horas semanais de prática	—	—	—	1,0	38,0	16,79	11,83

Curso de Especialização em Psicoterapia

O curso foi ministrado em regime b-learning, e possui uma carga horária formal de 552 horas, correspondendo à soma das 402 de componentes curriculares teóricas e 150 horas de supervisão clínica. O 1.º ano inclui os módulos *Fundamentos teóricos e investigação em psicoterapia I* (50 h), *Princípios de intervenção clínica I* (22 h), *Psicopatologia I* (22 h), *Prática deliberada em psicoterapia I* (22 h) e *Monitorização de processos e resultados* (18 h). O 2.º ano integra *Fundamentos teóricos e científicos em psicoterapia II* (30 h), *Formulação e vídeo-formação de casos clínicos* (34 h), *Psicopatologia II* (20 h), *Prática deliberada em psicoterapia II* (30 h) e *Psicoterapia para crianças e adolescentes I* (20 h). Por fim, o 3.º ano, não incluído na amostra deste estudo, contempla *Princípios de intervenção clínica II* (24 h), *Psicoterapia para crianças e adolescentes II* (26 h), *Formulação e vídeo-formação de casos clínicos II* (28 h), *Treino e prática experiencial* (24 h) e *Seminários temáticos* (32 h).

Instrumentos

Consentimento Informado e Questionário Sociodemográfico

O consentimento informado (Anexo II) foi preenchido e devolvido pelos participantes por e-mail antes do início da recolha de dados. Após a sua receção, os participantes completaram um questionário sociodemográfico, disponibilizado através da plataforma *Google Forms* (Anexo III). Este questionário incluiu questões relativas à idade, sexo, anos de experiência profissional, horas semanais de prática clínica e modelo teórico seguido na prática psicoterapêutica.

Vídeos de Evocação Clínica

Os participantes acederam aos vídeos utilizados como estímulos através da plataforma online Skillsetter, uma ferramenta desenvolvida especificamente para cursos de psicoterapia assentes nos princípios da prática deliberada. O conjunto incluía nove vídeos, cada um correspondendo a segmentos de uma sessão de psicoterapia simulada por atores e concebida para representar desafios clínicos frequentes e relevantes (Allen et al., 2023; Anderson et al., 2019; Anderson & Patterson, 2013). Este tipo de simulação tem sido amplamente utilizado como estratégia eficaz para treinar competências terapêuticas, permitindo a evocação de processos clínicos num contexto controlado (Hill & Lent, 2006). Quando surgiam momentos críticos na interação terapêutica, o vídeo era interrompido e pedia-se ao participante que produzisse uma resposta dirigida ao paciente, adotando a perspetiva do terapeuta. Este procedimento está alinhado com metodologias contemporâneas de treino clínico baseadas em simulação e resposta ativa, frequentemente utilizadas na formação de psicoterapeutas (Boswell & Castonguay, 2007; Rousmaniere, 2017).

Os cenários incluíam os seguintes desafios expressos pelo paciente:

- Vídeo 1 (48 s): definição de objetivos para alcançar mudanças rápidas.
- Vídeo 2 (70 s): confusão sobre o processo terapêutico, com exigência de maior profissionalismo e celeridade na obtenção de resultados.
- Vídeo 3 (49 s): desmotivação e frustração face à perceção de ausência de progresso.
- Vídeo 4 (58 s): frustração associada à crença de que o psicoterapeuta não está a conseguir ajudar.
- Vídeo 5 (57 s): adoção de um conselho terapêutico sem melhorias, originando exigência e pedidos de explicação.

- Vídeo 6 (53 s): sensação de estar a ser avaliada pelo terapeuta.
- Vídeo 7 (32 s): questionamento direto sobre a impressão que o terapeuta tem sobre si.
- Vídeo 8 (115 s): dúvidas quanto à capacidade do terapeuta para apoiar e compreender.
- Vídeo 9 (144 s): percepção de progresso insuficiente face às expectativas.

Facility Interpersonal Skills Performance Task

A evolução das competências clínicas dos formandos foi avaliada através da Facilitative Interpersonal Skills Performance Task (Anderson & Patterson, 2013), um instrumento baseado em desempenho que visa medir a eficácia interpessoal do psicoterapeuta quando confrontado com situações terapêuticas particularmente desafiantes (Allen et al., 2023; Anderson et al., 2009; Anderson et al., 2019). Integra oito dimensões de competência interpessoal: fluência verbal, esperança e expectativas positivas, persuasão, expressão emocional, afeto/aceitação/compreensão, empatia, capacidade de criar aliança terapêutica e ruptura e reparação da aliança (Anderson & Patterson, 2013; Anderson et al., 2019).

Este modelo de treino e avaliação de competências inclui duas componentes principais: (a) uma tarefa de desempenho, na qual o participante observa uma vinheta em vídeo e produz uma resposta verbal como se estivesse a conduzir a sessão; e (b) um sistema de cotação subsequente, no qual avaliadores treinados classificam a qualidade dessa resposta com base nas dimensões definidas (ver Anexo IV; Anderson & Patterson, 2013; Anderson et al., 2019).

A cotação é realizada com uma escala de Likert de 5 pontos, aplicada a cada um dos oito itens. A escala varia entre 1 (“Pouco característico”) e 5 (“Altamente característico”), acompanhada de descritores operacionais que orientam a decisão dos avaliadores (Anderson & Patterson, 2013; Anderson et al., 2019). Pontuações de 1 ou 2 refletem um desempenho insuficiente, o ponto 3 corresponde a um nível adequado ou neutro, e pontuações de 4 ou 5 indicam respostas competentes ou altamente competentes, compatíveis com intervenções interpessoais eficazes (Anderson & Patterson, 2013).

A consistência interna dos itens revela valores moderados a elevados, com coeficientes α entre .43 e .89 (Anderson & Patterson, 2013). A FIS apresenta igualmente evidência de validade ecológica, uma vez que os vídeos representam interações terapêuticas realistas e emocionalmente exigentes, aproximando-se das condições habituais da prática

clínica (Anderson et al., 2009; de Jong et al., 2024). Além disso, a prática deliberada com estes cenários tem sido identificada como promotora de melhorias nas respostas interpessoais, reforçando a utilidade aplicada do instrumento (Anderson et al., 2019).

Facilitative Interpersonal Skills – Self Report (FIS-SR)

Para avaliar a percepção dos participantes acerca das suas próprias FIS, utilizou-se a FIS-SR (Anexo V), um questionário composto pelos mesmos oito itens que integram a FIS, mas adaptados para formato de autoavaliação (Anderson et al., 2019). A estrutura do instrumento reorganiza a sequência original dos itens, apresentando-os da seguinte forma: (1) Empatia, (2) Afeto, Aceitação e Compreensão, (3) Persuasão, (4) Capacidade de Criar Aliança, (5) Esperança e Expetativas Positivas, (6) Rutura e Reparação da Aliança, (7) Fluência Verbal e (8) Expressão Emocional. A cotação está invertida relativamente à FIS, variando de (1) “Concordo Muito” a (5) “Discordo Muito”. O instrumento demonstra uma consistência interna elevada, com $\alpha = .91$ (Anderson et al., 2019). A utilização desta prova permitiu comparar diretamente a autoavaliação com a heteroavaliação.

Procedimento

A recolha de dados referente ao primeiro ano da formação foi conduzida por Catalino (2024). Nesta fase, os participantes preencheram o consentimento informado e posteriormente o questionário sociodemográfico, este último recolhido de forma anónima. Após este procedimento inicial, foi-lhes concedido acesso à plataforma Skillsetter para responder às nove vinhetas em vídeo, produzindo uma resposta única gravada por webcam para cada cenário. Em simultâneo, receberam a FIS-SR em formato digital, que preencheram e devolveram eletronicamente. Este conjunto de procedimentos correspondeu ao momento pré seminários de formação. Após a conclusão do módulo *Prática Deliberada em Psicoterapia I* (fevereiro-maio de 2024), os participantes voltaram a responder às vinhetas no Skillsetter e completaram novamente a FIS-SR.

No presente estudo, foram recolhidos os dados relativos ao segundo ano da formação. Dado que o consentimento informado e o questionário sociodemográfico tinham sido completados no primeiro ano, estes instrumentos não foram reaplicados. Assim, os participantes realizaram apenas as respostas às vinhetas em vídeo na plataforma Skillsetter e

preencheram novamente a FIS-SR em formato digital, ambos exclusivamente após o módulo *Prática Deliberada em Psicoterapia II* (março–junho de 2025), correspondendo ao terceiro momento avaliativo.

Análise de Dados

Para a cotação das respostas, foram designados três avaliadores, todos estudantes do segundo ano do mestrado em Psicologia Clínica no ISPA. Antes do início do processo, os avaliadores realizaram um estudo aprofundado do manual da FIS, assegurando um domínio consistente dos critérios de cotação. Todas as respostas foram cotadas individualmente, prevenindo potenciais enviesamentos decorrentes da influência entre avaliadores. Numa fase inicial, cada avaliador cotou de forma independente o primeiro vídeo do terceiro momento avaliativo, seguindo-se uma reunião destinada à discussão das discrepâncias e ao alinhamento dos critérios interpretativos. Confirmada a uniformidade dos procedimentos, procedeu-se à cotação individual dos três vídeos subsequentes (vídeos 2, 3 e 4). Com estas quatro vinhetas cotadas, foi calculado CCI para avaliar a concordância interavaliadores, utilizando o software IBM SPSS Statistics 30.0.

O CCI relativo às medidas médias foi de .931, com intervalo de confiança entre .921 e .941, indicando uma confiabilidade excelente, de acordo com os critérios de Koo e Li (2016) e estatisticamente significativo ($p < .001$).

Após confirmada a confiabilidade interavaliadores, procedeu-se à cotação individual dos restantes vídeos. Os dados da FIS Performance Task e da FIS-SR foram introduzidos no IBM SPSS Statistics 30.0, utilizando-se, no caso da FIS, a média das pontuações atribuídas pelos três avaliadores em cada dimensão de cada vídeo-estímulo para efeitos de análise estatística.

Para permitir a comparabilidade entre instrumentos, as pontuações da FIS-SR, originalmente invertidas e organizadas numa ordem distinta, foram recodificadas e reordenadas no SPSS de forma a coincidirem com a direção e a estrutura de itens da FIS: (1) Fluência Verbal; (2) Esperança e Expetativas Positivas; (3) Persuasão; (4) Expressão Emocional; (5) Afeto, Aceitação e Compreensão; (6) Empatia; (7) Capacidade de Criar Aliança; (8) Rutura e Reparação da Aliança.

Dado que o estudo compreende três momentos avaliativos, todas as análises estatísticas incidiram sobre a comparação sistemática destes três tempos de avaliação (M1 = início do curso de especialização em psicoterapia; M2 = após a conclusão do primeiro módulo de formação em PD; M3 = após a conclusão do segundo módulo de formação em PD).

Com o objetivo de obter uma visão geral das mudanças ao longo do tempo, foi calculado um índice global para ambos os instrumentos (FIS global e FIS-SR global), correspondendo à média dos oito itens em cada um dos três momentos avaliativos.

Para os índices globais, o pressuposto de normalidade foi sempre cumprido. Assim, foram realizadas ANOVAs de medidas repetidas, incluindo uma ANOVA de um fator (Momento) para cada instrumento (FIS e FIS-SR) e uma ANOVA de dois fatores (Momento × Tipo de avaliação: auto vs. hetero), de modo a analisar potenciais interações. Sempre que necessário, aplicou-se o ajuste de Greenhouse–Geisser perante violações da esfericidade, considerando que este pressuposto é frequentemente violado em amostras pequenas e que tal violação pode aumentar o erro tipo I (Blanca et al., 2023; Haverkamp & Beauducel, 2017; Mauchly, 1940).

Nas análises item a item, os procedimentos estatísticos foram selecionados de acordo com o objetivo e o cumprimento dos pressupostos. No caso específico da análise de interações entre autoavaliação e heteroavaliação para cada item, verificou-se que o pressuposto de normalidade não era cumprido. Por esse motivo, a avaliação das interações item a item foi conduzida exclusivamente através de comparações não paramétricas, recorrendo ao teste de Wilcoxon para cada par de itens correspondentes (auto vs. hetero), com aplicação de correção de Bonferroni.

Nas restantes análises item a item, relativas à variação ao longo dos três momentos avaliativos, aplicaram-se ANOVAs de medidas repetidas sempre que o pressuposto de normalidade estava assegurado. Quando necessário, utilizou-se o ajuste de Greenhouse–Geisser devido a violações da esfericidade. Nos casos em que a normalidade não se verificou, recorreu-se aos testes de Friedman como alternativa não paramétrica. Quando identificada significância estatística, realizaram-se os testes pós-hoc apropriados: correção de Bonferroni nos pós-hoc das ANOVAs e testes de Wilcoxon com correção de Bonferroni após os Friedman.

A correção de Bonferroni resultou num nível de significância ajustado de $\alpha = .0167$ ($0.05/3$), o qual foi aplicado a todos os testes de Wilcoxon, tanto nas comparações entre auto e hetero como nos pós-hoc subsequentes aos testes de Friedman.

Resultados

Verificação do Pressuposto de Normalidade (FIS e FIS-SR)

A normalidade foi avaliada através do teste de Shapiro–Wilk aplicado aos resíduos dos modelos de medidas repetidas, e encontra-se apresentada nos Anexos VII e VIII.

Tanto a pontuação global da FIS como da FIS-SR evidenciaram distribuição aproximadamente normal nos três momentos avaliados (FIS: $p = .093, .186$ e $.094$; FIS-SR: $p = .578, .982$ e $.195$), cumprindo o pressuposto de normalidade.

Ao nível dos itens da FIS, verificou-se normalidade em todos os momentos para I1 e I4 ($p > .05$). Em contraste, os Itens I2, I3, I5, I6, I7 e I8 revelaram violações de normalidade em pelo menos um dos momentos (por exemplo, I2: $p = .008, .004$ e $.047$; I3: $p = .006$; I5: $p < .001$; I6: $p = .035$; I7: $p < .001$; I8: $p = .002$).

Relativamente à FIS-SR, os Itens I2, I3, e I8 apresentaram resíduos normalmente distribuídos em todos os momentos ($p > .05$). Por outro lado, observaram-se violações de normalidade nos Itens I1, I4, I5, I6 e I7, cada um apresentando pelo menos um momento com valores de p inferiores a $.05$ (por exemplo, I1: $p = .001$; I5: $p = .017$; I6: $p = .045$; I7: $p = .009$).

FIS Global (Heteroavaliação)

Foi realizada uma ANOVA de medidas repetidas (Anexo IX) para comparar a pontuação global da FIS, nos três momentos de avaliação. A Tabela 2 apresenta as estatísticas descritivas (médias e desvios-padrão) e os resultados principais da ANOVA de medidas repetidas, enquanto a Tabela 3 apresenta os resultados das comparações pós-hoc (Bonferroni). A Figura 1 ilustra graficamente a evolução das médias da FIS global ao longo dos três momentos.

Os resultados revelaram diferenças estatisticamente significativas nas pontuações médias globais da FIS ao longo do tempo, $F(1.19, 19.02) = 72.73, p < .001, \eta^2_p = .82$. De

acordo com os critérios propostos por Cohen (1988), este valor de η^2_p indica um tamanho de efeito elevado, sugerindo que cerca de 82 % da variância nas pontuações da FIS se deve ao momento de avaliação.

As comparações pós-hoc (Bonferroni) revelaram diferenças significativas entre todos os momentos: M1 < M2 ($p < .001$); M2 < M3 ($p < .001$); e M1 < M3 ($p < .001$). As médias aumentaram sistematicamente entre os três momentos, indicando melhorias consistentes no desempenho avaliado.

Tabela 2

Estatísticas Descritivas e Resultados Principais da ANOVA de Medidas Repetidas Para Pontuação Global da FIS.

	M	DP	$F(df_1, df_2)$	p	η^2_p
M1	2.68	0.31			
M2	2.96	0.21			
M3	3.40	0.24			
ANOVA (Correção de Greenhouse–Geisser)	—	—	$F(1.19,$ $19.02) =$ 72.73	$< .001$.82

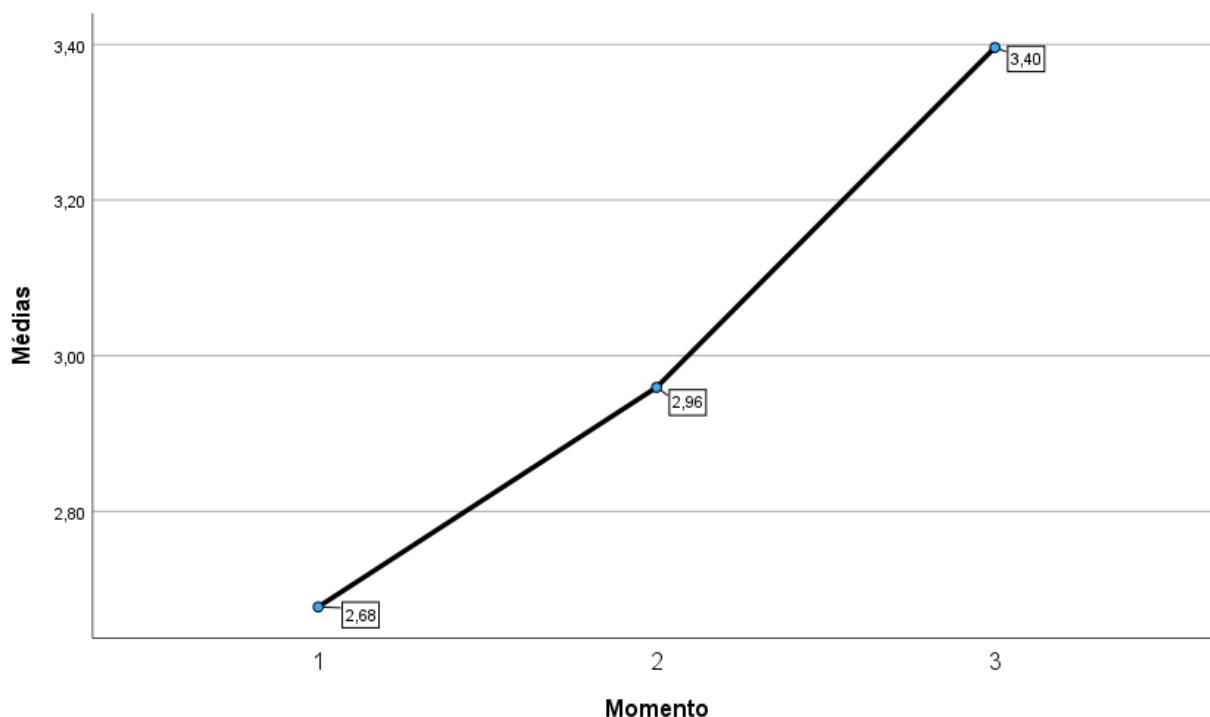
Tabela 3

Comparações pairwise (Bonferroni) para as médias da FIS Global nos três momentos de avaliação

	Diferença média (I–J)	Erro padrão	p (Bonferroni)	IC _{95%} da diferença	Interpretação
M1 – M2	–0.282	0.041	$< .001$	[–0.39, –0.17]	M1 < M2
M2 – M3	–0.437	0.051	$< .001$	[–0.57, –0.30]	M2 < M3
M1 – M3	–0.719	0.081	$< .001$	[–0.94, –0.50]	M1 < M3

Figura 1

Evolução das médias da FIS global ao longo dos três momentos de avaliação



Comparações da FIS (heteroavaliação) Entre Momentos (Itens 1 a 8)

Cada barra da Figura 2 representa a média das cotações dos três avaliadores em cada momento de avaliação (M1, M2 e M3), para cada um dos oito itens analisados. Observa-se que, após o treino em PD, os participantes evidenciaram melhorias significativas em todas as competências avaliadas, refletindo um aumento geral das pontuações ao longo do tempo. Em todos os itens foram verificadas diferenças estatisticamente significativas entre os momentos.

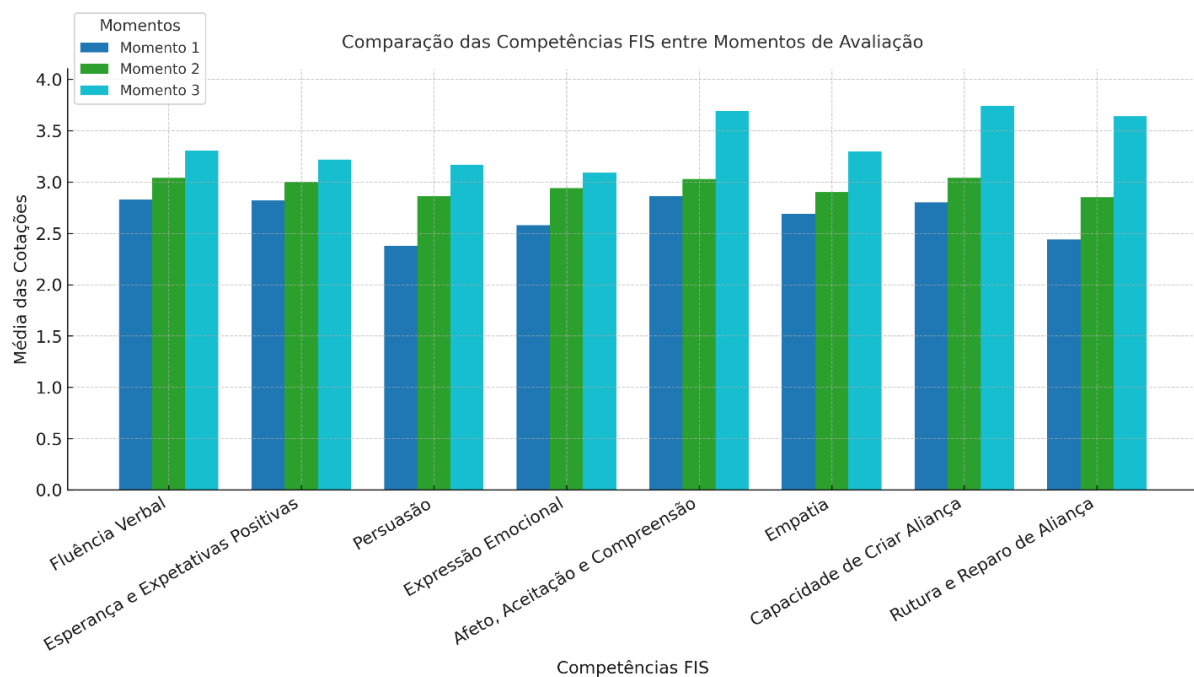
A análise dos oito itens da FIS revelou melhorias significativas ao longo dos três momentos de avaliação. Para os Itens 1 – *Fluência Verbal* e 4 – *Expressão Emocional* analisados por ANOVA de medidas repetidas (Anexo X), observaram-se efeitos estatisticamente significativos do Momento: Item 1 $F(1,35, 21,55) = 37,24, p < .001, \eta^2_p = .70$; Item 4 $F(2, 32) = 41,02, p < .001, \eta^2_p = .72$. As comparações pós-hoc com ajuste de Bonferroni mostraram diferenças significativas entre todos os momentos em ambos os itens: no Item 1, entre M1–M2 ($p < .001$), M1–M3 ($p < .001$) e M2–M3 ($p = .002$); e no Item 4, entre M1–M2

($p < .001$), M1–M3 ($p < .001$) e M2–M3 ($p = .029$), evidenciando um aumento sistemático dos valores médios do M1 para o M3.

Para os restantes itens (2, 3 e 5 a 8), analisados através do teste de Friedman (Anexo XI), verificaram-se efeitos estatisticamente significativos do Momento em todos os casos, com valores de qui-quadrado compreendidos entre $\chi^2(2) = 27.67$ e $\chi^2(2) = 34.00$ (todos $p < .001$). As comparações pós-hoc de Wilcoxon (Anexo XII), utilizando correção de Bonferroni para três comparações (α corrigido = .0167), evidenciaram diferenças estatisticamente significativas entre todos os pares de momentos (M1–M2, M1–M3 e M2–M3) em todos os itens. Os valores dos testes ($Z = -2.87$ a -3.62 , $p \leq .004$) indicam aumentos sistemáticos dos valores médios entre momentos consecutivos, refletindo melhorias consistentes ao longo do tempo.

Figura 2

Comparação das médias das competências FIS ao longo dos três momentos de avaliação



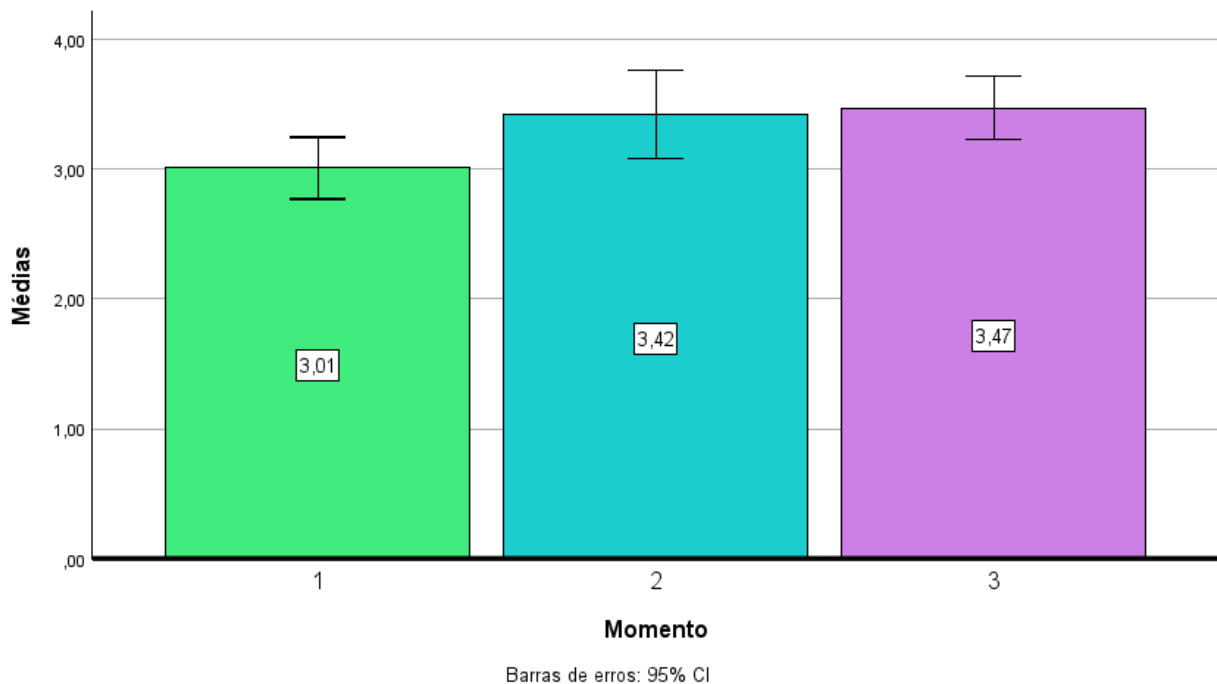
Evolução da FIS-SR (Autoavaliação) ao Longo dos Três Momentos

A ANOVA de medidas repetidas (Anexo XIII) revelou um efeito estatisticamente significativo do Momento da Avaliação na pontuação global da escala e tamanho de efeito elevado (Cohen, 1988), $F(2, 32) = 4.84$, $p = .015$, $\eta^2_p = .23$. As comparações pós-hoc com correção de Bonferroni indicaram uma diferença significativa entre o Momento 1 e o

Momento 3 (diferença média = -0.46 , $p = .005$), não tendo sido observadas diferenças significativas entre o Momento 1 e o Momento 2 (diferença média = -0.41 , $p = .098$) nem entre o Momento 2 e o Momento 3 (diferença média = -0.05 , $p = 1.000$). Os valores médios correspondentes aos três momentos encontram-se representados na Figura 3.

Figura 3

Médias da Pontuação global da FIS-SR nos Três Momentos de Avaliação



A análise longitudinal item a item revelou padrões distintos de evolução ao longo dos três momentos. Nos itens analisados com ANOVA de medidas repetidas (Anexo XIV), o Item 2 - *Esperança e Expetativas Positivas* não apresentou diferenças significativas entre momentos, $F(2, 32) = 2,80$, $p = .076$, $\eta^2_p = .15$. Em contraste, observaram-se efeitos estatisticamente significativos do Momento no Item 3 - *Persuasão*, $F(2, 32) = 4,43$, $p = .020$, $\eta^2_p = .22$, e no Item 8 - *Rutura e Reparo da Aliança*, $F(2, 32) = 4,21$, $p = .024$, $\eta^2_p = .21$. As comparações pós-hoc com correção de Bonferroni indicaram, em ambos os itens, diferenças estatisticamente significativas entre M1 e M3 (Item 3: $p = .022$; Item 8: $p = .006$), não se verificando entre M1 e M2 (Item 3: $p = .090$; Item 8: $p = .607$) nem entre M2 e M3 (Item 3: $p = 1.000$; Item 8: $p = .589$).

Relativamente aos itens analisados com o teste de Friedman (Anexo XV), observaram-se diferenças estatisticamente significativas para todos os itens: Item 1 – *Fluência Verbal*, $\chi^2(2) = 9,41$, $p = .009$; Item 4 - *Expressão Emocional*, $\chi^2(2) = 8,89$, $p = .012$; Item 5 - *Afeto, Aceitação e Compreensão*, $\chi^2(2) = 8,33$, $p = .016$; Item 6 - *Empatia*, $\chi^2(2) = 7,39$, $p = .025$; Item 7 - *Capacidade de Criar Aliança*, $\chi^2(2) = 12,32$, $p = .002$. Foram realizados testes de Wilcoxon (Anexo XVI) com correção de Bonferroni para múltiplas comparações ($\alpha_{\text{corrigido}} = .05/3 = .0167$). Observaram-se diferenças estatisticamente significativas entre M1 e M3 nos Itens 1 ($p = .009$), 4 ($p = .014$), 5 ($p = .007$) e 6 ($p = .014$), bem como entre M1 e M2 no Item 6 ($p = .015$). No Item 7, apenas a comparação M1–M2 foi estatisticamente significativa ($p = .009$).

Comparação da FIS (heteroavaliação) e FIS-SR (autoavaliação) ao Longo dos Três Momentos Avaliativos

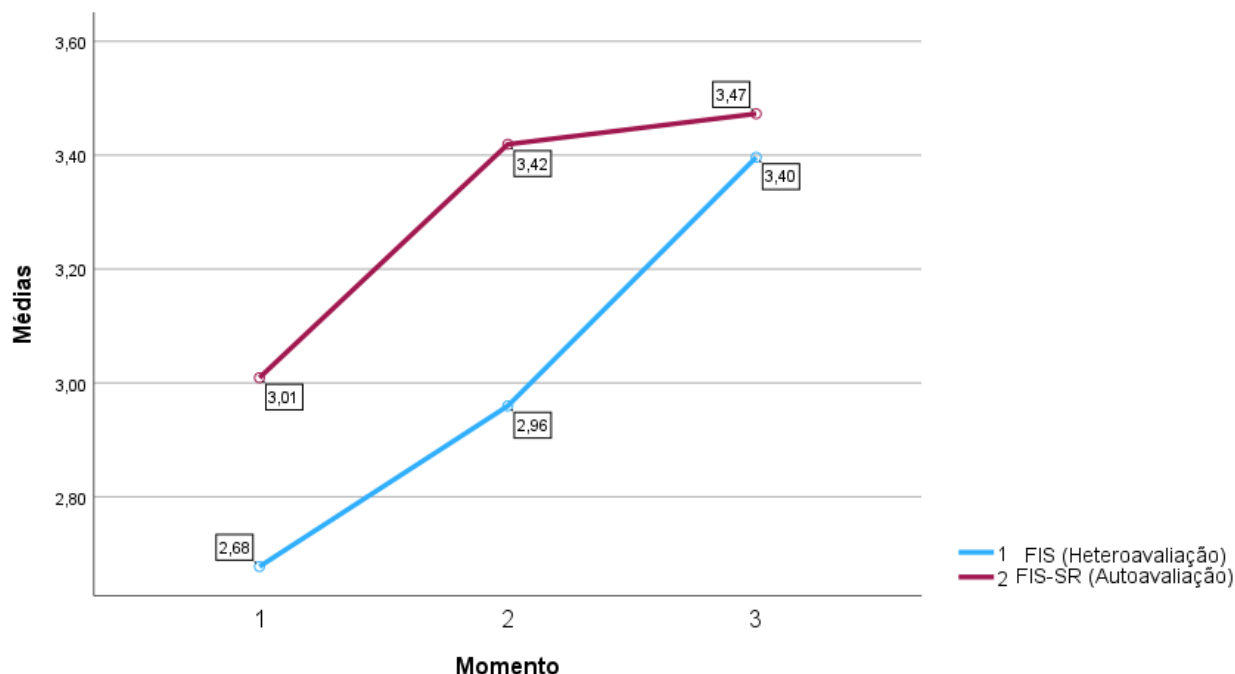
A observação do gráfico da Figura 3 sugere uma diminuição progressiva das distâncias ao longo dos três momentos, indicando que os valores parecem aproximar-se com o passar do tempo. Embora esta tendência visual seja evidente, o teste estatístico global não atingiu significância ($p = .058$), ficando, porém, muito próximo do valor de referência para significância estatística.

Na comparação Global entre os dois tipos de avaliação a ANOVA de medidas repetidas de 2 Fatores (Anexo XVII) revelou um efeito significativo de Momento, $F(2, 32) = 19.59$, $p < .001$, $\eta^2_p = .55$, indicando variação ao longo do tempo, e um efeito significativo do Tipo de Avaliação, $F(1, 16) = 11.43$, $p = .004$, $\eta^2_p = .42$, evidenciando que um tipo de avaliação apresentou valores consistentemente superiores ao outro. Crucialmente, a interação Momento \times Tipo de Avaliação não foi significativa, $F(2, 32) = 3.13$, $p = .058$, $\eta^2_p = .16$, indicando que a diferença entre os tipos de avaliação se manteve estável ao longo dos três momentos, sem evidência estatística de aproximação ou afastamento.

As comparações *pairwise* revelam que a FIS (heteroavaliação) apresentou valores significativamente inferiores à FIS-SR (autoavaliação) (diferença média = -0.289 , SE = 0.086 , $p = .004$).

Figura 4

Evolução das Médias Globais da FIS (heteroavaliação) e da FIS-SR (autoavaliação) nos Três Momentos Avaliativos



Aplicou-se o teste de Wilcoxon para comparar os resultados da FIS (heteroavaliação) e da FIS-SR (autoavaliação) nos Itens 1 a 8 ao longo de três momentos (Anexo XVIII). Dado que, para cada item, foram realizadas três comparações, utilizou-se a correção de Bonferroni, estabelecendo-se $\alpha = .0167$. Os resultados das comparações encontram-se apresentados na Tabela 4 e descrevem-se de seguida.

Observaram-se padrões distintos de significância estatística nos itens da escala. Os Itens 1 – *Fluência Verbal*, 2 - *Esperança e Expetativas Positivas* e 3 - *Persuasão* não evidenciaram diferenças estatisticamente significativas após a correção. No Item 1, nenhum dos momentos apresentou resultados e estatisticamente significativos (M1: $Z = -0.83$, $p = .407$; M2: $Z = -0.99$, $p = .320$). O Item 2 seguiu o mesmo padrão (M1: $Z = -1.01$, $p = .313$; M2: $Z = -0.92$, $p = .356$; M3: $Z = -0.40$, $p = .687$), e de igual modo o Item 3 não apresentou variações estatisticamente significativas (M1: $Z = -1.25$, $p = .210$; M2: $Z = -0.85$, $p = .394$; M3: $Z = -1.23$, $p = .218$).

Nos restantes itens, verificaram-se efeitos limitados a determinados momentos. No Item 8 - *Rutura e Reparo da Aliança*, apenas M1 apresentou diferenças estatisticamente significativas ($Z = -3.53, p < .001$), enquanto M2 ($Z = -2.32, p = .020$) e M3 ($Z = -0.85, p = .394$) não atingiram o limiar ajustado. No Item 6 - *Empatia*, somente M2 evidenciou diferenças ($Z = -2.56, p = .011$), não se verificando significância em M1 ($Z = -2.37, p = .018$) nem em M3 ($Z = -2.07, p = .039$).

Por fim, alguns itens apresentaram diferenças consistentes em dois momentos. No Item 5 - *Afeto, Aceitação e Compreensão*, observaram-se diferenças em M1 ($Z = -2.58, p = .010$) e M2 ($Z = -2.72, p = .006$), não se verificando diferenças em M3 ($Z = -0.36, p = .717$). O mesmo padrão ocorreu no Item 7 - *Capacidade de Criar Aliança*, com significância em M1 ($Z = -2.77, p = .006$) e M2 ($Z = -2.58, p = .010$), enquanto M3 não evidenciou diferenças ($Z = -1.52, p = .129$), adicionalmente, no Item 4 - *Expressão Emocional*, M1 ($Z = -3.20, p = .001$) e M2 ($Z = -2.66, p = .008$) mantiveram-se estatisticamente significativos após a correção; M3 ($Z = -2.07, p = .039$) não permaneceu significativo.

De forma geral, verificou-se que as diferenças entre autoavaliação e heteroavaliação foram mais frequentes em M1 e M2, sendo menos evidentes em M3, momento em que vários itens deixaram de apresentar valores estatisticamente significativos.

Tabela 4

Resultados do teste de Wilcoxon para a comparação entre heteroavaliação (FIS) e autoavaliação (FIS-SR) nos três momentos

Item	Momento	Z	p	Sig. Bonferroni
1	M1	-0.83	.407	
	M2	-0.99	.320	
	M3	-2.11	.035	
2	M1	-1.01	.313	

	M2	-0.92	.356	
	M3	-0.40	.687	
3	M1	-1.25	.210	
	M2	-0.85	.394	
	M3	-1.23	.218	
4	M1	-3.20	.001	*
	M2	-2.66	.008	*
	M3	-2.07	.039	
5	M1	-2.58	.010	*
	M2	-2.72	.006	*
	M3	-0.36	.717	
6	M1	-2.37	.018	
	M2	-2.56	.011	*
	M3	-2.07	.039	
7	M1	-2.77	.006	*
	M2	-2.58	.010	*
	M3	-1.52	.129	
8	M1	-3.53	< .001	*
	M2	-2.32	.020	
	M3	-0.85	.394	

Nota. Resultados estatisticamente significativos após correção de Bonferroni para $\alpha = .0167$ assinalados com *.

Discussão

Os resultados obtidos sugerem que os formandos evidenciam melhorias objetivas e estatisticamente significativas nas competências avaliadas externamente ao longo da formação (Tabela 3), o que dá suporte à hipótese (H1) e está em linha com a literatura que documenta os efeitos positivos da prática deliberada estruturada no desenvolvimento de competências clínicas (Brand et al., 2025; Lagerberg et al., 2024; Larsson et al., 2023; Mahon, 2024; Olsson et al., 2024).

Adicionalmente, os resultados evidenciam uma melhoria sistemática e consistente em todos os itens avaliados externamente ao longo dos três momentos (Figura 2), evidenciando um progresso claro do desempenho dos formandos. Enquanto no M1 e M2 as médias se situavam abaixo (M1) ou muito próximas (M2) do ponto neutro (“3”), indicando desempenho inicial ainda marcado por défices ou competências em desenvolvimento, no M3 as médias ultrapassam esse nível, aproximando-se de indicadores de mestria (“4–5”). Este padrão é inteiramente compatível com evidência empírica que demonstra que a prática deliberada continuada, promove ganhos graduais, robustos e acumulativos em competências complexas ao longo do tempo (Ericsson & Harwell, 2019; Lagerberg et al., 2024; Larsson et al., 2023; Olsson et al., 2024; Mahon, 2024).

A evolução positiva na heteroavaliação contrasta com o padrão sistemático de autoavaliações mais elevadas em todos os momentos (Figura 3), o que vai ao encontro da hipótese (H2) e replica de forma consistente o que tem sido documentado na literatura empírica. Estudos que utilizam especificamente o método FIS apontam que os formandos tendem a atribuir a si próprios classificações superiores às dos avaliadores externos, mesmo quando o desempenho objetivo melhora ao longo do treino (Longley et al., 2023). Esta discrepância é igualmente observada noutras investigações sobre desenvolvimento de competências clínicas, onde as autoavaliações se revelam tipicamente inflacionadas face às avaliações externas, sobretudo em competências interpessoais e relacionais (Banner et al., 2025; Bennett-Levy, 2019; Parker & Waller, 2015; Walfish et al., 2012). Assim, o padrão encontrado no presente estudo não só vai ao encontro da hipótese (H2), como se alinha com

evidência robusta que demonstra que, em fases formativas, a percepção subjetiva de competência tende a exceder a competência observável medida externamente.

A análise das discrepâncias entre FIS e FIS-SR (Tabela 4) revela padrões diferenciados entre itens, que se enquadram nos modelos teóricos sobre autoavaliação e desenvolvimento de competências clínicas. Nos Itens 1 (Fluência Verbal), 2 (Esperança/Expectativas) e 3 (Persuasão) não emergiram discrepâncias estatisticamente significativas, o que poderá refletir competências mais observáveis e comportamentais, cujo desempenho tende a ser percebido de forma relativamente convergente entre avaliadores internos e externos. Este padrão é consistente com evidência que sugere maior precisão da autoavaliação em tarefas com critérios claros e baixa ambiguidade avaliativa (Eva & Regehr, 2005; Sitzmann et al., 2010).

Em contraste, no Item 8 (Rutura e Reparação da Aliança) surgiu discrepância apenas em M1, um resultado plausível dado a complexidade da monitorização interna de ruturas e estratégias de reparação, especialmente em fases iniciais de formação (Safran et al., 2011; Eubanks et al., 2018). O Item 6 (Empatia) apresentou discrepância apenas em M2, possivelmente refletindo oscilações temporárias na confiança subjetiva em competências relacionais centrais, um fenómeno documentado em terapeutas em formação (Longley et al., 2023).

Por fim, os Itens 4 (Expressão Emocional), 5 (Afeto/Aceitação/Compreensão) e 7 (Aliança) mostram discrepâncias presentes em M1–M2, mas não em M3, sugerindo que competências de natureza humanística são particularmente vulneráveis à sobreavaliação inicial. Este padrão é consistente com a evidência de que competências relacionais e empáticas tendem a ser as mais sobreavaliadas pelos formandos, devido à sua forte componente subjetiva (Longley et al., 2023), e a redução dessas discrepâncias até M3 é igualmente compatível com literatura que tem sugerido que o treino estruturado e o feedback continuado promovem uma maior calibragem entre autoavaliação e heteroavaliação ao longo da formação (Anderson et al., 2019).

No que respeita à hipótese exploratória (H3), observa-se uma diminuição progressiva da discrepância entre autoavaliação e heteroavaliação ao longo do tempo. No entanto, esta tendência não atingiu significância estatística, o que indica que a convergência entre as duas medidas deve ser interpretada com cautela. A aproximação observada resulta sobretudo do aumento expressivo das pontuações atribuídas pelos avaliadores externos, ainda

que as autoavaliações também evidenciem uma melhoria estatisticamente significativa entre M1 e M3. Assim, embora exista um movimento visual de convergência (Figura 4), este parece refletir principalmente a evolução objetiva captada pela heteroavaliação, mais do que um processo robusto de calibragem metacognitiva.

Este padrão é, ainda assim, coerente com modelos teóricos que descrevem um alinhamento gradual entre percepção subjetiva e desempenho real à medida que a experiência se consolida e a autorregulação se torna mais precisa (Bandura, 1997). Ao mesmo tempo, enquadra-se na literatura que evidencia que as autoavaliações tendem a apresentar limitações sistemáticas, sobretudo em tarefas complexas e interpessoais, sendo frequentemente menos sensíveis a mudanças reais de desempenho (Eva & Regehr, 2005). O desaparecimento das discrepâncias nos Itens 4, 5, 7 e 8 é compatível com esta tendência, apesar de a hipótese exploratória (H3) não ter sido suportada estatisticamente, sugerindo apenas um indício inicial de calibragem metacognitiva ao longo da formação.

Limitações

O presente estudo apresenta várias limitações que devem ser consideradas na interpretação dos resultados. Uma primeira limitação diz respeito ao tamanho reduzido da amostra, que restringe o poder estatístico, especialmente para detetar efeitos de interação, os quais tipicamente exigem amostras maiores para serem identificados com fiabilidade (Bakeman, 2005; Gelman & Hill, 2007; Maxwell & Delaney, 2004). Outra limitação relaciona-se com a utilização da correção de Bonferroni que, embora útil para reduzir o risco de erro tipo I, aumenta substancialmente o risco de erro tipo II, dificultando a deteção de efeitos reais de menor magnitude (Perneger, 1998; Nakagawa, 2004).

Adicionalmente, embora a FIS e a FIS-SR sejam instrumentos robustos, a autoavaliação permanece vulnerável a enviesamentos cognitivos e limitações metacognitivas bem documentadas, sobretudo em fases iniciais da formação (Eva & Regehr, 2005; Dunlosky & Metcalfe, 2009). No presente estudo, esta vulnerabilidade pode ter sido acentuada pelo facto de alguns participantes terem completado várias tarefas de forma consecutiva. A realização prolongada e contínua de atividades avaliativas, especialmente sem pausas estruturadas, está associada a maior carga cognitiva e a uma diminuição da precisão metacognitiva, conduzindo a julgamentos menos calibrados (Dunlosky & Metcalfe, 2009). Assim, não é possível excluir a hipótese de que parte da discrepância entre autoavaliação e

heteroavaliação reflita, em parte, efeitos de fadiga ou redução momentânea de autorregulação durante a resposta. Outra limitação relaciona-se com o desenho temporal das medições. Embora exista uma medida de referência no início do primeiro ano, não foi recolhida uma linha de base equivalente no início do segundo ano, antes do módulo de “Prática Deliberada em Psicoterapia II”. Assim, torna-se difícil isolar a contribuição específica deste módulo, uma vez que entre M2 e M3 decorreram também outras atividades formativas que podem ter influenciado o desenvolvimento de competências.

A estrutura da FIS-SR pode constituir uma limitação adicional, por ser composta exclusivamente por itens invertidos, um formato associado a maior carga cognitiva, potenciais ambiguidades interpretativas e artefactos psicométricos (Marsh, 1996; DiStefano & Motl, 2006; van Sonderen et al., 2013). Acresce que a FIS-SR não foi preenchida imediatamente após cada resposta às vinhetas, introduzindo um desfasamento temporal suscetível de gerar enviesamentos de memória ou reconstrução (Dunlosky & Metcalfe, 2009). Uma forma simples de mitigar esta possível limitação no futuro seria recorrer ao SkillSetter (ou plataforma equivalente) para apresentar os itens na ordem correta e imediatamente após cada vídeo, reduzindo a carga cognitiva e evitando que os participantes tenham de reconstruir mentalmente o desempenho antes de responder.

Por último, a cotação do terceiro momento foi realizada por uma equipa distinta daquela que avaliou os dois primeiros momentos, o que implica a possibilidade de pequenas variações nos critérios aplicados ou na interpretação das escalas. A variabilidade entre avaliadores é uma limitação reconhecida em investigação com medidas comportamentais, dado que diferenças subtis nos procedimentos, formação ou estilo de avaliação podem introduzir discrepâncias sistemáticas nas classificações (McGraw & Wong, 1996). Além disso, mesmo em contextos de formação estruturada, os avaliadores tendem a divergir na forma como aplicam e ponderam os mesmos critérios, o que reforça a inevitabilidade deste tipo de variação (Paunov et al., 2024).

Implicações Para Investigação Futura

A investigação futura deverá recorrer a amostras maiores e delineamentos mais controlados, de modo a aumentar o poder estatístico e permitir análises mais sensíveis do desenvolvimento de competências ao longo do treino. Será igualmente importante

padronizar as condições de resposta, evitando que os participantes completem todas as tarefas de forma consecutiva ou sob fadiga, dado o impacto conhecido na precisão da autoavaliação.

Além disso, recomenda-se a recolha de linhas de base claramente definidas em cada ano de formação e o controlo de módulos formativos paralelos, de modo a clarificar a contribuição específica de cada intervenção. Futuras investigações deverão também examinar propriedades alternativas da FIS-SR, incluindo versões com itens não totalmente invertidos e administrações imediatas após cada tarefa, reduzindo potenciais efeitos de memória e carga cognitiva.

Por fim, futuras investigações poderiam examinar se a participação dos psicólogos em psicoterapia pessoal contribui para o desenvolvimento de competências interpessoais facilitadoras, dada a evidência de que esta experiência está associada a maior autoconsciência, empatia e responsividade clínica (Macran & Shapiro, 1998; Norcross & Guy, 2005; Orlinsky et al., 2011).

Conclusão

O presente estudo contribui para a investigação emergente sobre prática deliberada na formação em psicoterapia, evidenciando que este método promove melhorias significativas nas competências interpessoais facilitadoras dos formandos, corroborando a literatura existente. A discrepância observada entre autoavaliação e heteroavaliação, bem como a sua redução parcial ao longo do tempo, salienta a importância de processos de treino que favoreçam uma perceção mais calibrada das próprias competências. Apesar das limitações identificadas, os resultados reforçam a relevância do treino sistemático das FIS como componente essencial na formação de terapeutas. Futuras investigações deverão recorrer a delineamentos mais robustos e explorar variáveis adicionais potencialmente relevantes para o desenvolvimento destas competências.

Referências

- Allen, J. J., Parker, A., & Ogles, B. M. (2023). A review of the facilitative interpersonal skills performance task and rating method. *Clinical Psychology: Science and Practice*, 31(2), 211–223. <https://doi.org/10.1037/cps0000187>
- Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of clinical psychology*, 65(7), 755-768. <https://doi.org/10.1002/jclp.20583>
- Anderson, T., & Patterson, C. (2013). *Facilitative interpersonal skill task and rating method*. Ohio University.
- Anderson, T., McClintock, A. S., Himawan, L., Song, X., & Patterson, C. L. (2016). A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *Journal of Consulting and Clinical Psychology*, 84(1), 57–66. <https://doi.org/10.1037/ccp0000060>
- Anderson, T., Perlman, M. R., McCarrick, S. M., & McClintock, A. S. (2019). Modeling therapist responses with structured practice enhances facilitative interpersonal skills. *Journal of Clinical Psychology*, 76(4), 659- 675. <https://doi.org/10.1002/jclp.22911>
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3), 379–384. <https://doi.org/10.3758/BF03192707>
- Baldwin, S. A., & Imel, Z. E. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 258–297). John Wiley & Sons.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman
- Banner, S. E., Rock, A. J., Cosh, S. M., Schutte, N., & Rice, K. (2025). Self-reflection on competence: Metacognitive process and barriers of self-assessment in psychologists. *Advances in Health Sciences Education*, 30, 1–21. <https://doi.org/10.1007/s10459-025-10434-7>
- Bennett-Levy, J. (2019). Why therapists should walk the talk: The theoretical and empirical case for personal practice in therapist training and professional development. *Journal of Behavior Therapy and Experimental Psychiatry*, 62, 133–145. <https://doi.org/10.1016/j.jbtep.2018.08.004>
- Blanca, M. J., Arnau, J., Alarcón, R., Bono, R., & Bendayan, R. (2023). Repeated measures ANOVA and adjusted F-tests when sphericity is violated: Which procedure is best? *Frontiers in Psychology*, 14, 1192453. <https://doi.org/10.3389/fpsyg.2023.1192453>

- Bohart, A. C., & Tallman, K. (2010). Clients: The neglected common factor in psychotherapy. In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 83–111). American Psychological Association.
- Boswell, J. F., & Castonguay, L. G. (2007). Psychotherapy training: Suggestions for core ingredients and future research. *Psychotherapy: Theory, Research, Practice, Training*, 44(4), 378–383. <https://doi.org/10.1037/0033-3204.44.4.378>
- Brand, J., Miller-Bottome, M., Vaz, A., & Rousmaniere, T. (2025). Deliberate practice supervision in action: The Sentio supervision model. *Journal of Clinical Psychology*, 81(6), 462–472. <https://doi.org/10.1002/jclp.23790>
- Catalino, D. L. S. (2024). *A eficácia da prática deliberada na melhoria de competências interpessoais facilitadoras e a sua autoavaliação* [Dissertação de mestrado, ISPA – Instituto Universitário]. Repositório ISPA. <http://hdl.handle.net/10400.12/10228>
- Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy*, 52(3), 337–345. <https://doi.org/10.1037/pst0000015>
- Clements-Hickman, A. L., & Harris, K. R. (2024). Deliberate practice for psychotherapy skills: Recommendations and implications based on the state of the science. *Professional Psychology: Research and Practice*, 55(6), 563–572. <https://doi.org/10.1037/pro0000592>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^a ed.). Lawrence Erlbaum Associates.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- de Jong, K., Conijn, J. M., Gallagher, R. A., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, 85, 1-19. <https://doi.org/10.1016/j.cpr.2021.102002>
- Diamond, G., Wlodek, B., Arthey, S., & Parker, S. (2025). A systematic review of deliberate practice in psychotherapy: Definitions, operationalization, and preliminary outcomes. *Psychotherapy*, 62(2), 113–131. <https://doi.org/10.1037/pst0000569>
- DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*, 13(3), 440–464. https://doi.org/10.1207/s15328007sem1303_6

- Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. SAGE Publications.
- Ericsson, A., & Pool, R. (2016). *Peak: Secrets from the new science of expertise*. Houghton Mifflin Harcourt.
- Ericsson, K. A., & Harwell, K. W. (2019). Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future research. *Frontiers in Psychology, 10*, 2396. <https://doi.org/10.3389/fpsyg.2019.02396>
- Eubanks, C. F., Muran, J. C., & Safran, J. D. (2018). Alliance rupture repair: A meta-analysis. *Psychotherapy, 55*(4), 508–519. <https://doi.org/10.1037/pst0000185>
- Eva, K. W., & Regehr, G. (2005). Self-assessment in the health professions: A reformulation and research agenda. *Academic Medicine, 80*(10 Suppl), S46–S54. <https://doi.org/10.1097/00001888-200510001-00015>
- Frank, J. D., & Frank, J. B. (1991). *Persuasion and healing: A comprehensive study of psychotherapy* (3rd ed.). Johns Hopkins University Press.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T., & Wampold, B. E. (2016). Do psychotherapists improve with time and experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counseling Psychology, 63*(1), 1–11. <https://doi.org/10.1037/cou0000131>
- Haverkamp, N., & Beauducel, A. (2017). Violation of the sphericity assumption and its effect on Type-I error rates in repeated measures ANOVA and multi-level linear models. *Frontiers in Psychology, 8*, 1841. <https://doi.org/10.3389/fpsyg.2017.01841>
- Heinonen, E., & Nissen-Lie, H. A. (2020). The professional and personal characteristics of effective psychotherapists: A systematic review. *Psychotherapy Research, 30*(4), 417–432. <https://doi.org/10.1080/10503307.2019.1620366>
- Hill, C. E., & Lent, R. W. (2006). A narrative and meta-analytic review of helping skills training: Time to revive a dormant area of inquiry. *Psychotherapy: Theory, Research, Practice, Training, 43*(2), 154–172. <https://doi.org/10.1037/0033-3204.43.2.154>
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy, 48*(1), 9–16. <https://doi.org/10.1037/a0022186>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine, 15*(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>

- Lagerberg, H., Boswell, J. F., Constantino, M. J., Andersson, G., & Carlbring, P. (2024). Does practice make perfect? The effects of an eight-week manualized deliberate practice course with peer feedback on patient-rated working alliance in adults: A pilot randomized controlled trial. *Clinical Psychology in Europe*, 6(3), 1–19. <https://doi.org/10.32872/cpe.12353>
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (2nd ed., pp. 203–223). Oxford University Press.
- Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, 50(1), 42–51. <https://doi.org/10.1037/a0030682>
- Larsson, J., Werthén, D., Carlsson, J., Salim, O., Davidsson, E., Vaz, A., Sousa, D., & Norberg, J. (2023). Does deliberate practice surpass didactic training in learning empathy skills? – A randomized controlled study. *Nordic Psychology*, 77(1), 39–52. <https://doi.org/10.1080/19012276.2023.2247572>
- Laska, K. M., Gurman, A. S., & Wampold, B. E. (2013). Expanding the lens of evidence-based practice in psychotherapy: A common factors perspective. *Psychotherapy*, 51(4), 467–481. <https://doi.org/10.1037/a0034332>
- Longley, M., Kästner, D., Daubmann, A., Hirschmeier, C., Strauß, B., & Gumz, A. (2023). Prospective psychotherapists' bias and accuracy in assessing their own facilitative interpersonal skills. *Psychotherapy*, 60(4), 1–11. <https://doi.org/10.1037/pst0000506>
- Macdonald, J., & Mellor-Clark, J. (2015). Helpers' self-assessment biases before and after helping skills training. *The European Journal of Counselling Psychology*, 3(1), 1–17. <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01377/full>
- Macran, S., & Shapiro, D. A. (1998). The role of personal therapy for therapists: A review. *British Journal of Medical Psychology*, 71(1), 13–25. <https://doi.org/10.1111/j.2044-8341.1998.tb01364.x>
- Mahon, D. (2022). A scoping review of deliberate practice in the acquisition of therapeutic skills and practices. *Counselling and Psychotherapy Research*, 23(4), 965–981. <https://doi.org/10.1002/capr.12601>
- Mahon, D. (2024). “Don't just magically assume you're going to get better as a therapist as time goes on”: An exploration of the use of deliberate practice. *Counselling and Psychotherapy Research*, 24(3), 913–924. <https://doi.org/10.1002/capr.12746>
- Maroco, P. (2021). *Análise estatística com o SPSS Statistics* (8ª ed.). ReportNumber.

- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 70(4), 810–819. <https://doi.org/10.1037/0022-3514.70.4.810>
- Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics*, 11(2), 204–209. <https://doi.org/10.1214/aoms/1177731915>
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30–46. <https://doi.org/10.1037/1082-989X.1.1.30>
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Psychology Press.
- Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology*, 15(6), 1044–1045. <https://doi.org/10.1093/beheco/arh107>
- Nissen-Lie, H. A., Rønnestad, M. H., Høglend, P. A., Havik, O. E., Solbakken, O. A., Stiles, T. C., & Monsen, J. T. (2025). Deliberate practice in psychotherapy supervision: A promising paradigm for developing therapist competence. *Journal of Clinical Psychology*. Advance online publication. <https://doi.org/10.1002/jclp.70043>
- Norcross, J. C., & Guy, J. D. (2005). Personal therapy for psychotherapists. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 459–478). Oxford University Press.
- Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy*, 55(4), 303–315. <https://doi.org/10.1037/pst0000193>
- Nurse, K., O’Shea, M., & Sheen, J. (2024). Qualitative accounts from deliberate practice training for empathic communication among psychotherapy trainees. *Counselling and Psychotherapy Research*, 00, 1–12. <https://doi.org/10.1002/capr.12761>
- Olsson, M., Gustafsson, M., Carlsson, J., Rousmaniere, T., Bergbom, S., & Norberg, J. (2024). A pilot study of a deliberate practice intervention for therapist trainees. *Nordic Psychology*, 77(2), 165–184. <https://doi.org/10.1080/19012276.2024.2313532>
- Orlinsky, D. E., Schofield, M. J., Schroder, T., & Norcross, J. C. (2011). Psychotherapists’ lifelong learning. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-based therapist contributions* (2nd ed., pp. 319–344). Oxford University Press.

- Parker, J., & Waller, G. (2015). Therapists' self-assessment of competence in cognitive-behavioural therapy: Differences between novice and experienced clinicians. *Behaviour Research and Therapy*, *64*, 85–90.
<https://doi.org/10.1016/j.brat.2014.11.007>
- Paunov, T., Weck, F., Heinze, P. E., Maaß, U., & Kühne, F. (2024). Competence ratings in psychotherapy training – A complex matter. *Cognitive Therapy and Research*, *48*(3), 500–510. <https://doi.org/10.1007/s10608-023-10445-x>
- Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ*, *316*(7139), 1236–1238. <https://doi.org/10.1136/bmj.316.7139.1236>
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, *21*(2), 95–103.
<https://doi.org/10.1037/h0045357>
- Rousmaniere, T. (2017). *Deliberate practice for psychotherapists: A guide to improving clinical effectiveness*. Routledge.
- Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. In *Psychotherapy relationships that work: Evidence-based responsiveness*, 2nd ed. (pp. 224–238). <https://doi.org/10.1093/acprof:oso/9780199737208.003.0011>
- Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education*, *9*(2), 169–191. <https://doi.org/10.5465/amle.9.2.zqr169>
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32*(9), 752–760. <https://doi.org/10.1037/0003-066X.32.9.752>
- Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist*, *69*(3), 218–229.
<https://doi.org/10.1037/a0035099>
- Truax, C. B., & Carkhuff, R. R. (1967). *Toward effective counseling and psychotherapy: Training and practice*. Aldine.
- van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLOS ONE*, *8*(7), e68967.
<https://doi.org/10.1371/journal.pone.0068967>
- Vaz, A., Sousa, D., & Afonseca, M. (2023). *Treino em psicoterapia. Como usar a prática deliberada para melhorar resultados clínicos*. Pactor.

- Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*(2), 639–644. <https://doi.org/10.2466/02.07.17.PR0.110.2.639-644>
- Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World psychiatry, 14*(3), 270-277. <https://doi.org/10.1002/wps.20238>
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.

Anexos

Anexo I

Revisão da Literatura

I.1 — Eficácia da Psicoterapia

A discussão contemporânea sobre a eficácia da psicoterapia tem origem nos trabalhos pioneiros de Hans Eysenck, cuja revisão de 1952 sugeria que um maior envolvimento em tratamento psicoterapêutico estava associado a menores taxas de recuperação em perturbações neuróticas, chegando mesmo a levantar a possibilidade de que a intervenção pudesse ser prejudicial para alguns indivíduos (Eysenck, 1952). Esta conclusão, amplamente divulgada à época, moldou de forma significativa a percepção pública e profissional sobre os efeitos da psicoterapia durante várias décadas (Smith & Glass, 1977).

A contestação a esta posição emergiu rapidamente. Investigadores como Rosenzweig (1954), Bergin (1963) e Strupp (1963) salientaram falhas metodológicas substanciais na revisão de Eysenck e defenderam que a psicoterapia, quando adequadamente estudada, demonstra benefícios clínicos relevantes. A viragem decisiva ocorreu com a primeira meta-análise formal sobre a eficácia da psicoterapia (Smith & Glass, 1977), que indicou que os indivíduos submetidos a intervenções psicoterapêuticas obtinham resultados superiores aos de grupos sem tratamento em aproximadamente 75% dos casos.

A literatura acumulada desde então consolidou um consenso robusto de que a psicoterapia é, de forma consistente, eficaz para uma grande diversidade de perturbações psicológicas (Abu Sabra et al., 2023; Barkham & Lambert, 2021; Barkowski et al., 2020; Cuijpers, 2017; Eilsen et al., 2022; Hilbert et al., 2019; Lambert, 2013; Leichsenring et al., 2022; Munder et al., 2019; Rahmani et al., 2020; Sousa, 2017; Weber et al., 2021). Diversas meta-análises recentes apontam que os efeitos da psicoterapia se equiparam, na maioria das condições, aos das intervenções farmacológicas (Cuijpers et al., 2020; Høstmælingen et al., 2021; Kappelmann et al., 2020). Além disso, há evidências de que os ganhos terapêuticos provenientes da psicoterapia tendem a ser mais duradouros ao longo do tempo, quando comparados com os obtidos exclusivamente pela medicação (Cuijpers et al., 2013; Imel et al., 2008; Wampold, 2007). Em muitos casos, a combinação entre psicoterapia e farmacoterapia potencia ainda mais a melhoria clínica (Cuijpers et al., 2020; Høstmælingen et al., 2021).

À medida que a evidência empírica foi consolidando a ideia de que a psicoterapia produz efeitos positivos robustos, a investigação passou a centrar-se num problema distinto:

determinar em que condições esses efeitos variam. As sínteses teóricas e meta-analíticas mais recentes têm mostrado que, de modo geral, os ganhos obtidos com a intervenção psicológica tendem a manter-se relativamente estáveis independentemente da perturbação alvo ou da orientação conceptual que estrutura o tratamento (Sousa, 2017). Desde o trabalho seminal de Smith e Glass (1977), tem sido amplamente argumentado que não existem diferenças estatisticamente significativas entre abordagens psicoterapêuticas distintas, um princípio que continua a ser suportado por meta-análises contemporâneas (Abu Sabra et al., 2023; Cuijpers, 2017; Cuijpers et al., 2021; Eijsen et al., 2022). De forma geral, estima-se que cerca de 67% dos clientes em psicoterapia apresentem melhoras clinicamente significativas, contrastando com cerca de 33% em grupos sem tratamento (Lambert, 2013).

Este padrão verifica-se numa ampla gama de perturbações: depressão (Cuijpers, 2017; Cuijpers et al., 2021; Eijsen et al., 2022; Munder et al., 2019); perturbações de ansiedade (Barkowski et al., 2020; Rahmani et al., 2020); perturbação obsessivo-compulsiva (Abu Sabra et al., 2023); e perturbações do comportamento alimentar (Hilbert et al., 2019).

A evidência acumulada ao longo de mais de sete décadas permite, assim, afirmar que a psicoterapia é um tratamento eficaz, seguro e transversalmente aplicável a diferentes quadros psicopatológicos, mantendo ainda vantagens claras em termos de durabilidade dos resultados e adaptabilidade a diversas necessidades clínicas.

I.2 — Para Além da Eficácia: Os Desafios da Psicoterapia

Embora a investigação documente de forma consistente a eficácia da psicoterapia, é igualmente necessário reconhecer um conjunto de desafios atuais que condicionam a prática clínica (Sousa, 2017). Entre estes destaca-se o problema das taxas de abandono, amplamente reportado em diversas meta-análises, que apontam valores aproximados de 20% (Linardon et al., 2019) e até cerca de 28% (Dixon & Linardon, 2020; Iliakis et al., 2021). Evidência acumulada reforça esta tendência, indicando que um em cada cinco clientes termina a terapia prematuramente (Swift & Greenberg, 2012; Swift & Greenberg, 2014), sendo que este abandono ocorre sobretudo nas sessões iniciais (Fernandez et al., 2015). A aliança terapêutica tem-se revelado um fator determinante para explicar este fenómeno (Ferreira et al., 2023; Linardon et al., 2019).

A literatura evidencia também que uma proporção considerável de clientes não apresenta melhorias com a psicoterapia (cerca de 30%), e que aproximadamente 10% sofre

deterioração clínica (Barkham & Lambert, 2021). Estes resultados acompanham conclusões de estudos anteriores, que estimam que entre 5–10% dos pacientes pioram e que 35–40% não registam progressos, mesmo quando acompanhados por terapeutas altamente treinados e supervisionados (Lambert et al., 2005; Shimokawa et al., 2010). Em contexto infanto-juvenil, a taxa de deterioração pode ser ainda maior, situando-se entre 10 % e 20 % (Warren et al., 2010).

Outro desafio prende-se com os efeitos negativos da psicoterapia (Klatte et al., 2023), um conceito que inclui eventos adversos relacionados ou não com os procedimentos terapêuticos e que contribuem para o agravamento do estado clínico (Moritz et al., 2019). A prevalência destes efeitos varia substancialmente entre estudos: 52,6% (Moritz et al., 2019), 56,6% (Strauss et al., 2021) e 70,5% (Abeling et al., 2018). Uma parte significativa destes eventos emerge no âmbito da relação terapêutica (Safran et al., 2005). Assim, a consciência do terapeuta sobre a possibilidade de deterioração torna-se indispensável para prevenir danos (Klatte et al., 2023).

Um problema adicional diz respeito à limitada capacidade dos terapeutas de identificarem clientes em deterioração (Hatfield et al., 2010; Lambert, 2013). No estudo de Hannan et al. (2005), apenas 3 em 550 clientes foram reconhecidos pelos terapeutas como estando a piorar (0,01%), quando efetivamente 40 se encontravam em deterioração (7,3%). Existe igualmente discrepância entre as perceções do terapeuta e do cliente relativamente à qualidade da aliança terapêutica (Hartmann et al., 2014), com uma correlação média de apenas .36 (Tryon et al., 2007). Estudos recentes reforçam esta tendência, sugerindo que a experiência subjetiva do cliente frequentemente diverge da do terapeuta, influenciando a responsividade e o risco de abandono (Gjerstad et al., 2024). Outros estudos apontam para diferenças relevantes na importância atribuída a componentes específicos da terapia, contribuindo para falhas de responsividade (Bachelor, 2013; Timulak, 2010).

Adicionalmente, verifica-se uma tendência marcada para que os terapeutas sobrestimem a própria eficácia (Lambert, 2013). Evidências mostram que 25% se consideram entre os 10% melhores profissionais e nenhum se avalia abaixo da média (Walfish et al., 2012).

A implementação de sistemas de monitorização contínua dos resultados e de feedback tem-se revelado eficaz para mitigar estes desafios (Barkham et al., 2023; de Jong et al., 2021; Lambert et al., 2018), contribuindo para melhorias terapêuticas na ordem dos 8% (Barkham et al., 2023). Estes sistemas reduzem a deterioração (Lambert et al., 2018), promovem a

diminuição da sintomatologia e podem reduzir o abandono terapêutico (de Jong et al., 2021). Além disso, permitem reforçar a colaboração, ajustar intervenções e fortalecer a aliança (Norcross & Lambert, 2018). Por fim, a literatura destaca a necessidade de o terapeuta operar a partir de um modelo teórico sólido, mas simultaneamente flexível, de modo a adaptar a intervenção às singularidades de cada cliente e construir um racional terapêutico claro e coerente (Norcross & Wampold, 2018).

I.3 — Fatores Comuns: Fundamentos e Integração no Modelo Contextual

A investigação sobre os mecanismos de mudança em psicoterapia tem demonstrado que diferentes abordagens partilham elementos fundamentais que explicam, em grande medida, a sua eficácia. Esta ideia remonta a Rosenzweig (1936), cujo veredicto do pássaro Dodô sugeria que diversas terapias obtêm resultados semelhantes por integrarem princípios comuns, como a relação terapêutica, a catarse emocional, a personalidade do terapeuta e a criação de um enquadramento coerente.

Autores posteriores aprofundaram esta perspetiva. Frank e Frank (1991) propuseram quatro fatores universais envolvidos na mudança: uma relação emocionalmente significativa, um contexto seguro, um racional explicativo e a participação em rituais terapêuticos coerentes. A consistência destes efeitos ao longo de diferentes orientações reforça a ideia de que os mecanismos centrais da mudança são amplamente partilhados (Bohart & Tallman, 2010).

Nesta linha, Norcross e Lambert (2019) estimam que os fatores extraterapêuticos representam cerca de 40% dos resultados, os fatores comuns 30%, as técnicas específicas 15% e as expectativas 15%. Embora aproximativos, estes valores salientam a natureza interdependente destes elementos (Cuijpers et al., 2019; Lambert & Barley, 2001). Entre os fatores comuns, a aliança terapêutica destaca-se como um dos preditores mais robustos. Definida pelo acordo nas tarefas, objetivos e vínculo emocional (Bordin, 1979; Sousa, 2017). A aliança está consistentemente associada à melhoria clínica, independentemente da abordagem (Horvath & Symonds, 1991; Horvath et al., 2011). Outros fatores relacionais relevantes incluem empatia, aceitação, colaboração e a capacidade de reparação de ruturas (Elliott et al., 2011; Farber et al., 2022; Safran & Kraus, 2014; Safran et al., 2011; Tryon & Winograd, 2011). Lambert (2013) agrupou estes contributos em fatores de suporte,

aprendizagem e ação, evidenciando que os fatores comuns operam dinamicamente ao longo do processo terapêutico.

A integração destes contributos teóricos é consolidada no modelo contextual de Wampold que contrasta com o modelo médico ao defender que a eficácia terapêutica resulta sobretudo de fatores relacionais e contextuais (Wampold, 2001; Wampold & Imel, 2015). Segundo o modelo, a mudança ocorre através de três vias: a relação real, a criação de expectativas positivas e os ingredientes específicos, entendidos não como técnicas isoladas, mas como ações coerentes com o racional terapêutico que reforçam a esperança (Wampold & Imel, 2015).

A relação real, marcada por genuinidade, empatia e intencionalidade (Gelso, 2014; Wampold, 2015), distingue-se de uma relação social comum pela profundidade emocional e confidencialidade. As expectativas terapêuticas emergem quando o paciente adota um racional que substitui crenças desmoralizantes (Wampold, 2015). Antes destas vias se ativarem, é crucial o estabelecimento de confiança inicial, etapa na qual ocorrem mais abandonos (Fernandez et al., 2021; van Dijk et al., 2023).

De forma global, a literatura indica que as variáveis com maior poder preditivo dos resultados incluem aliança, empatia, expectativas, colaboração, feedback, aceitação positiva, genuinidade e congruência (Norcross & Lambert, 2018; Wampold, 2015).

O modelo contextual integra também os efeitos do terapeuta, isto é, as diferenças individuais que influenciam significativamente os resultados e a adaptação às características do paciente, incluindo cultura, preferências e valores, aumentando a credibilidade do processo terapêutico e fortalecendo a aliança (Norcross & Lambert, 2018; Wampold, 2015).

I.4 — Efeitos do Terapeuta na Eficácia Psicoterapêutica

A investigação em psicoterapia tem-se focado amplamente nos modelos de intervenção e nos efeitos específicos das técnicas, atribuindo menor atenção aos fatores comuns e, em particular, ao impacto do terapeuta (Barkham et al., 2017; Norcross & Lambert, 2018). Contudo, há evidência crescente de que os terapeutas diferem sistematicamente na eficácia alcançada com os seus clientes, contribuindo de forma significativa para os resultados clínicos (Delgado et al., 2022; Heinonen & Nissen-Lie, 2020; Johns et al., 2019; Wampold, 2015).

Os “efeitos do terapeuta” referem-se à variância dos resultados explicada exclusivamente pelo profissional, estimada entre 5% e 8%, embora com ampla heterogeneidade entre estudos (Barkham et al., 2017; Chow et al., 2015; Johns et al., 2019). Estes valores não refletem qualidade absoluta, mas sim diferenças relativas entre terapeutas. Se estes efeitos fossem nulos, os clientes obteriam resultados equivalentes independentemente de quem os atendesse (Johns et al., 2019). A literatura indica que alguns terapeutas obtêm resultados sistematicamente superiores aos dos seus pares, independentemente do modelo teórico ou do tipo de perturbação (Miller et al., 2013; Wampold & Brown, 2005). Este padrão é consistente com evidência adicional que atribui parte significativa dos resultados a características individuais dos terapeutas (Baldwin & Imel, 2013; Laska et al., 2013) e ao papel central de competências relacionais como empatia, autenticidade e capacidade responsiva (Rogers, 1957; Truax & Carkhuff, 1967).

Além disso, investigações recentes sugerem que, mesmo quando atuam em condições semelhantes e utilizam métodos comparáveis, alguns terapeutas alcançam resultados clinicamente superiores aos dos seus colegas, um fenômeno amplamente documentado ao longo de várias décadas (Barkham et al., 2017; Heinonen & Nissen-Lie, 2020; Johns et al., 2019; Pereira et al., 2023; Vaz et al., 2023; Wampold, 2015).

Apesar de frequentemente assumidas como relevantes, variáveis demográficas e formativas como idade, gênero, anos de experiência, formação adicional ou adesão a manuais terapêuticos apresentam uma capacidade preditiva reduzida dos resultados (Anderson et al., 2009; Brown et al., 2005; Owen & Hilsenroth, 2014; Spengler & Pilipis, 2015). Em vez disso, competências relacionais, como a capacidade de construir uma aliança terapêutica sólida, reparar rupturas e adaptar-se responsivamente ao cliente, surgem como preditores robustos da eficácia (Baldwin et al., 2007; Safran & Muran, 2000; Wampold & Imel, 2015). As Facilitative Interpersonal Skills (FIS) explicam igualmente parte substancial desta variabilidade (Anderson et al., 2015, 2016).

Evidências recentes indicam que a monitorização sistemática do progresso clínico pode reduzir significativamente a variabilidade entre terapeutas. Uma meta-análise envolvendo mais de 4.500 clientes e 131 terapeutas apontou para que o uso consistente de feedback de progresso diminuiu as diferenças nos resultados clínicos entre profissionais mais e menos eficazes (Delgado et al., 2022).

Processos desenvolvimentais contínuos, como a utilização sistemática de feedback do cliente e a capacidade de manter uma “insegurança profissional saudável”, parecem igualmente contribuir para trajetórias de melhoria ao longo da carreira (Lambert & Shimokawa, 2011; Nissen-Lie et al., 2015).

I.5 — Prática Deliberada

A prática deliberada tem sido amplamente estudada como um dos mecanismos mais robustos na explicação do desenvolvimento de expertise em diversos domínios de desempenho especializado. A partir dos trabalhos de Ericsson e colaboradores, consolidou-se a ideia de que o desempenho excepcional não resulta de talento inato, mas sim de um investimento prolongado em treino altamente estruturado, orientado para a melhoria contínua e sustentado por feedback rigoroso (Ericsson & Pool, 2016; Ericsson, 2018; Vaz et al., 2023). O estudo clássico de Ericsson e colaboradores, no início da década de 1990, demonstrou que violinistas de elite com melhor desempenho acumulavam significativamente mais horas de treino específico e desafiante (Ericsson et al., 1993). Padrões semelhantes foram observados noutras áreas: pianistas experientes mantinham níveis elevados de desempenho ao longo da idade quando continuavam a envolver-se em treino deliberado (Krampe & Ericsson, 1996); músicos profissionais apresentavam ganhos consistentes resultantes de treino orientado para melhorias graduais (Platz et al., 2014); e jogadores de xadrez altamente qualificados acumulavam várias vezes mais horas de treino específico do que pares menos proficientes (Charness et al., 2005). Em contexto clínico, programas formativos baseados em prática deliberada demonstraram igualmente uma eficácia superior na aquisição de competências médicas e cirúrgicas quando comparados com modelos tradicionais de formação (Hashimoto et al., 2015; Hunt et al., 2014; McGaghie et al., 2011).

Embora algumas revisões críticas no domínio do desporto sugeriram que a prática deliberada não explica isoladamente todas as diferenças de desempenho (Macnamara, 2014; Macnamara et al., 2016), análises posteriores evidenciaram que muitos desses estudos não utilizavam definições rigorosas do conceito, reforçando que a prática deliberada permanece um dos métodos mais sólidos para o desenvolvimento de competências complexas (Ericsson & Harwell, 2019; Miller et al., 2020).

A prática deliberada caracteriza-se por envolver atividades estruturadas com o objetivo explícito de melhorar o desempenho através da definição de metas concretas, da

repetição sistemática de micro-habilidades e da obtenção de feedback imediato, idealmente fornecido por um especialista que ajusta o treino ao nível de competência do aprendiz (Clements-Hickman & Reese, 2020; Ericsson et al., 1993; Miller et al., 2020; Rousmaniere et al., 2017). Este processo exige esforço cognitivo elevado, períodos prolongados de concentração e uma tolerância significativa ao desconforto inerente a trabalhar de forma consistente fora da zona de conforto (Ericsson & Pool, 2016; Hill et al., 2020). A sua eficácia depende do cumprimento rigoroso destes critérios, distinguindo-se de outras formas de prática, como a *naive practice*, baseada apenas na repetição, a *purposeful practice*, que não inclui supervisão especializada, ou a *structured practice*, conduzida de forma genérica e não individualizada (Ericsson & Pool, 2016; Ericsson & Harwell, 2019).

O mecanismo central que explica os ganhos obtidos através da prática deliberada reside no desenvolvimento e refinamento de representações mentais, isto é, estruturas cognitivas que organizam a informação relevante para a tarefa e permitem executar padrões complexos com maior precisão e economia de recursos (Ericsson, 2018, 2021). Com o treino repetido, estas representações tornam-se mais diferenciadas e funcionam como mapas internos que permitem avaliar o desempenho atual, identificar discrepâncias e orientar ajustamentos imediatos, tal como demonstrado em estudos com músicos e atletas (Ericsson & Harwell, 2019; Ericsson, 2020). Este processo é cumulativo e prolonga-se ao longo de anos, sendo mantido mesmo após a aquisição de níveis elevados de competência, contrastando com a ideia popular das “10.000 horas”, frequentemente mal interpretada (Ericsson et al., 2018). A investigação sugere que não existe um número fixo aplicável a todos os domínios, variando a quantidade de treino necessária entre áreas e entre indivíduos (Ericsson & Pool, 2016; Rousmaniere et al., 2017). Evidência neurocientífica reforça esta perspetiva ao indicar que o treino intensivo gera alterações específicas nas redes neurais associadas às competências praticadas, sublinhando a natureza altamente contextual da aprendizagem avançada (Ericsson & Pool, 2016).

Por fim, a prática deliberada diferencia-se de abordagens formativas centradas predominantemente na aprendizagem conceptual, como a leitura, as aulas teóricas ou a observação passiva, pois privilegia a aprendizagem procedimental, em que o foco recai no desenvolvimento direto de habilidades através de ciclos contínuos de execução, análise e correção (Mahon, 2022; Vaz & Rousmaniere, 2022).

I.6 — Expertise em Psicoterapia: O Papel da Prática Deliberada

A variabilidade consistente no desempenho entre terapeutas tem motivado um crescente interesse pelos processos subjacentes ao desenvolvimento da expertise clínica: o que distingue terapeutas consistentemente mais eficazes?

Estudos qualitativos recentes contribuíram significativamente para esta compreensão. A análise de Caçador et al. (2023) identifica três competências nucleares: a capacidade de promover autonomia e colaboração, a habilidade para facilitar uma exploração emocional profunda e significativa, e a criação de um ambiente relacional seguro, sustentado por autenticidade e presença empática. Os autores sublinham ainda a importância das sessões iniciais e da revisão sistemática de gravações como elementos fundamentais na melhoria contínua.

Em contraste, a maioria dos terapeutas tende a apresentar uma percepção inflacionada do próprio desempenho. Investigações comparando autoavaliações com avaliações externas apontam para discrepâncias sistemáticas, especialmente no domínio das competências interpessoais facilitadoras (Longley et al., 2023). Este padrão é coerente com evidência mais ampla que documenta limitações da reflexão espontânea e da metacognição clínica (Bennett-Levy, 2019; Macdonald & Mellor-Clark, 2015). Estudos qualitativos sugerem que avaliações mais precisas dependem da existência de modelos claros de competência e de contextos estruturados de treino (Banner et al., 2025). Adicionalmente, terapeutas mais eficazes parecem distinguir-se por uma capacidade mais acurada de monitorizar o impacto das suas intervenções e ajustar o processo terapêutico em função desse feedback (Ericsson & Pool, 2016; Vaz et al., 2023).

A literatura sobre expertise demonstra de forma consistente que a acumulação de anos de prática, por si só, não assegura progressão, o aperfeiçoamento avançado depende de métodos de treino sistemáticos, com objetivos claros, critérios explícitos e feedback frequente, exatamente os elementos que caracterizam a prática deliberada (Ericsson & Pool, 2016; Tracey et al., 2014). Quando estes princípios são incorporados na formação, observa-se uma diminuição da discrepância entre a percepção subjetiva e o desempenho efetivo (Longley et al., 2023; Macdonald & Mellor-Clark, 2015).

Embora inicialmente concebida para domínios com regras de desempenho bem definidas, a prática deliberada (PD) tem sido progressivamente adaptada à psicoterapia, evidenciando utilidade crescente (Clements-Hickman & Reese, 2020; Ericsson & Harwell,

2019). Estudos experimentais e quase-experimentais sugerem que programas estruturados segundo estes princípios promovem melhorias objetivas em competências centrais da relação terapêutica, incluindo empatia, responsividade, gestão de resistência e comunicação clínica, em diferentes populações de terapeutas (Axelsson et al., 2023; Barrett-Naylor et al., 2020; Clements-Hickman & Harris, 2024; Lagerberg et al., 2024; Larsson et al., 2023; Mahon, 2024; Newman et al., 2022; Shukla et al., 2021; Westra et al., 2021; Yamin et al., 2023). Este conjunto de evidências sugere que estratégias de prática focada e progressiva permitem desenvolver dimensões complexas do desempenho clínico que raramente emergem espontaneamente.

De forma complementar, investigações longitudinais indicam que terapeutas que integram rotinas de monitorização e PD apresentam progressos consistentes ao longo da carreira (Chow et al., 2015; Goldberg et al., 2016). A supervisão desempenha aqui um papel determinante, quando orientada pelos princípios da PD, favorece a identificação de pontos cegos, o acesso a feedback imediato e a prática repetida com objetivos específicos, fatores associados ao refinamento de competências relacionais, técnicas e autorregulatórias (Axelsson et al., 2023; Mahon, 2022; Olsson et al., 2024; Young & Maack, 2021).

Ferramentas de treino como role-plays sucessivos, exercícios graduais e trabalho intensivo com gravações têm demonstrado utilidade particular, sobretudo quando integradas em sequências orientadas de execução, análise e ajustamento (Mahon, 2022; Vaz & Rousmaniere, 2022). Estes procedimentos mostram-se eficazes para melhorar a gestão de situações emocionalmente exigentes e promover intervenções mais empáticas e responsivas (Di Bartolomeo et al., 2021; Shukla et al., 2021; Westra et al., 2021).

No conjunto, a investigação acumulada identifica **claramente** a prática deliberada como uma via empiricamente sustentada para o desenvolvimento contínuo de competências terapêuticas, baseada em princípios de treino sistemático, repetição intencional, feedback qualificado e progressão gradual (Ericsson & Pool, 2016; Rousmaniere et al., 2017; Rousmaniere, 2019).

A relevância destes princípios torna-se particularmente clara quando aplicada ao treino das *Facilitative Interpersonal Skills* (FIS), um conjunto de capacidades relacionais observáveis que demonstram prever de forma consistente os resultados clínicos (Anderson et al., 2015). Além da sua importância preditiva, estas competências revelam-se altamente

desenvolvíveis através de métodos de treino estruturados, nomeadamente aqueles inspirados na lógica da prática deliberada (Anderson et al., 2016).

I.7 — *Facilitative Interpersonal Skills (FIS)*

Os terapeutas mais eficazes tendem a ser aqueles que mobilizam de forma consistente as suas capacidades interpessoais, respondendo de modo ajustado às necessidades dos pacientes e promovendo benefícios clínicos independentemente de variáveis como diagnóstico ou traços de personalidade (Heinonen & Nissen-Lie, 2020). Estas competências, frequentemente conceptualizadas como FIS, englobam dimensões como empatia, expressividade emocional, reparação de ruturas e a promoção de expectativas positivas (Anderson et al., 2009; Norcross & Lambert, 2018). A investigação tem indicado que níveis mais altos de FIS tendem a associar-se a melhores resultados terapêuticos, traduzindo-se em maiores taxas de melhoria para os pacientes (Allen et al., 2023; Anderson et al., 2009; Anderson & Patterson, 2013; Anderson et al., 2016; Anderson et al., 2019; Anderson & Perlman, 2022).

A importância destas competências tem impulsionado o desenvolvimento de métodos rigorosos que permitam avaliá-las de forma padronizada e ecologicamente válida. Entre estes instrumentos destaca-se a FIS Performance Task, que recorre a situações clínico-interpessoais simuladas para avaliar diretamente a capacidade do terapeuta de responder a desafios relacionais complexos, constituindo atualmente uma das medidas mais robustas da responsividade interpessoal (Anderson et al., 2019; Rousmaniere, 2017).

I.8 — *Facilitative Interpersonal Skills Performance Task*

A FIS Performance Task é um método de avaliação baseado em desempenho, desenvolvido para medir diretamente a resposta do terapeuta a desafios interpessoais complexos através de vinhetas audiovisuais de elevada validade ecológica (Anderson et al., 2019; Rousmaniere, 2017). Em vez de recorrer a autorrelatos, este procedimento apresenta vídeos interrompidos em momentos críticos, nos quais o participante deve responder como se estivesse em sessão (Allen et al., 2023; Anderson et al., 2019). Este formato permite observar a capacidade do terapeuta para responder a momentos de elevada exigência relacional, captando dimensões essenciais como fluência verbal, promoção de esperança e expectativas positivas, persuasão terapêutica, expressão emocional, aceitação, empatia, e a competência

para estabelecer, manter e reparar a aliança terapêutica (Anderson & Patterson, 2013; Anderson et al., 2019). Estas dimensões, largamente convergentes com as condições terapêuticas propostas por Rogers (1957), posicionam a FIS Performance Task como um dos indicadores mais sensíveis da responsividade interpessoal do terapeuta (Rousmaniere, 2017).

O método FIS é composto por duas etapas complementares: a tarefa de desempenho e o processo de cotação (Anderson et al., 2019). Na primeira, os participantes assistem vinhetas audiovisuais de curta duração, cada uma representando uma interação clínica encenada com elevada carga emocional ou relacional (Allen et al., 2023; Anderson et al., 2019). Em momentos críticos, o vídeo é interrompido e o participante deve responder como se estivesse em sessão; as respostas são gravadas e avaliadas posteriormente por especialistas devidamente formados, recorrendo a um manual para classificar o desempenho em oito itens, utilizando uma escala Likert de 1 “Pouco característico” a 5 “Altamente característico” (Anderson et al., 2019; Anderson & Patterson, 2013). Pontuações mais baixas refletem défices de competência, enquanto valores elevados indicam um nível de mestria (Anderson et al., 2009).

Do ponto de vista psicométrico, a FIS Performance Task apresenta elevada validade ecológica e facial, atribuída à natureza realista das vinhetas, bem como uma consistência interna robusta entre os oito itens avaliados (Allen et al., 2023; Anderson & Patterson, 2013). Estudos longitudinais e experimentais sugerem que as pontuações na FIS predizem de forma consistente os resultados clínicos, reforçando a relevância destas competências enquanto pilares transversais da eficácia terapêutica (Allen et al., 2023; Anderson et al., 2009, 2016).

Um aspeto particularmente relevante para a formação clínica diz respeito à discrepância entre autoavaliação e desempenho observado. Longley et al. (2023) indicam que os terapeutas tendem a sobrestimar as suas competências relacionais quando comparadas com avaliações independentes e que embora exista uma correlação moderada entre ambos os tipos de avaliação, sugerindo que quem se percebe como mais competente tende, de facto, a apresentar melhor desempenho, a magnitude da sobrestimação é substancial. Curiosamente, fatores como idade, género ou experiência prévia não se associam significativamente às estimativas subjetivas, ao passo que autoconfiança elevada parece amplificar esta discrepância (Longley et al., 2023).

Reconhecer esta tendência pode facilitar uma atitude mais recetiva ao feedback, sobretudo em terapeutas em formação.

O interesse crescente nas FIS enquanto preditores de resultados tem incentivado a sua integração em programas de treino e supervisão. A investigação demonstra que intervenções formativas centradas nas FIS, incluindo observação de respostas-modelo, prática repetida e feedback orientado, promovem melhorias mensuráveis nessas competências (Anderson et al., 2019). Estudos adicionais indicam que este tipo de treino ajuda os terapeutas a aprofundar o foco nas experiências e emoções dos clientes (Anderson et al., 2022) e que abordagens formativas baseadas na relação terapêutica e nas FIS superam métodos essencialmente didáticos na aquisição de competências clínicas (Perlman et al., 2020, 2023).

Neste sentido, a FIS Performance Task tem sido utilizada não apenas como instrumento de avaliação, mas também como ferramenta pedagógica útil para fornecer feedback individualizado e monitorizar o progresso ao longo da formação (Allen et al., 2023; Longley et al., 2023). A sua aplicação no contexto da formação surge, assim, alinhada com a evidência mais ampla sobre os fatores do terapeuta que influenciam os resultados clínicos (Barkham et al., 2017; Caçador et al., 2023; Edmondstone et al., 2023; Johns et al., 2019; Heinonen & Nissen-Lie, 2020).

Face à variabilidade interprofissional amplamente documentada, o desenvolvimento das FIS emerge como uma estratégia particularmente promissora para reduzir discrepâncias no desempenho clínico e promover níveis mais elevados de eficácia entre terapeutas (Allen et al., 2023; Anderson et al., 2009). Para que tal ocorra, vários autores sublinham a necessidade de criar contextos de aprendizagem que valorizem a tolerância ao erro, a exploração de limitações e a receção ativa de feedback, elementos fundamentais para o desenvolvimento profissional contínuo (Edmondstone et al., 2023; Heinonen & Nissen-Lie, 2020).

I.9 — Conclusão

A investigação sobre psicoterapia evoluiu de um período de ceticismo inicial (Eysenck, 1952) para um consenso robusto quanto à sua eficácia, amplamente demonstrada por diversas meta-análises (Cuijpers, 2017; Cuijpers et al., 2020). Uma vez estabelecida a eficácia geral, o debate científico deslocou-se para os mecanismos de mudança e para os fatores que explicam a variabilidade nos resultados entre terapeutas.

Entre estes, os fatores comuns como a aliança terapêutica, a empatia, as expectativas e a responsividade destacam-se como determinantes transversais das melhorias clínicas

(Norcross & Lambert, 2018; Wampold & Imel, 2015). A importância destes elementos relacionais converge com a literatura sobre os efeitos do terapeuta, que indica diferenças consistentes no desempenho clínico entre terapeutas que trabalham em condições comparáveis. Esta variabilidade sugere que características individuais e competências interpessoais desempenham um papel central nos resultados terapêuticos, sendo as *Facilitative Interpersonal Skills* (FIS) identificadas como preditoras particularmente robustas (Anderson et al., 2009; Anderson et al., 2016).

Paralelamente, a investigação sobre expertise clínica tem sugerido que os anos de experiência, por si só, não produzem melhorias sistemáticas. Em contraste, o desenvolvimento profissional exige métodos estruturados que incluam objetivos claros, prática focada e feedback qualificado, que são princípios centrais da prática deliberada (PD) (Ericsson & Pool, 2016; Rousmaniere et al., 2017). Aplicada ao contexto clínico, a PD tem demonstrado resultados promissores, com ganhos observáveis em competências relacionais essenciais através de estratégias como role-plays sucessivos, feedback imediato e análise de gravações (Axelsson et al., 2023; Mahon, 2022).

No domínio da avaliação formativa, a FIS Performance Task destaca-se por permitir uma medição padronizada da responsividade interpessoal em situações próximas da prática real, funcionando simultaneamente como instrumento de avaliação e como guia para intervenções de treino (Anderson et al., 2019). Adicionalmente, a discrepância frequentemente observada entre autoavaliação e heteroavaliação, com tendência para sobrestimação por parte dos formandos sublinha a necessidade de integrar feedback externo e medidas objetivas nos programas de formação (Longley et al., 2023).

Em conjunto, estes contributos sustentam a relevância de modelos de formação que sejam experiencialmente ricos, orientados por feedback e centrados no treino de competências relacionais validadas empiricamente. Neste enquadramento, o presente estudo propõe examinar se um programa de PD conduz a melhorias mensuráveis nas competências interpessoais facilitadoras e se influencia a discrepância entre autoavaliação e avaliação externa. Assim, formulam-se as seguintes hipóteses: (H1) a prática deliberada resultará numa melhoria estatisticamente significativa das competências interpessoais facilitadoras, avaliadas através da FIS; (H2) a autoavaliação das FIS manter-se-á sistematicamente superior à heteroavaliação; e (H3 — exploratória) poderá observar-se uma tendência de diminuição da discrepância entre autoavaliação e heteroavaliação ao longo do tempo de formação.

I.10 — Referências

- Abeling, B., Müller, A., Stephan, M., Pollmann, I., & de Zwaan, M. (2018). Negative Effekte von Psychotherapie: Häufigkeit und Korrelate in einer klinischen Stichprobe. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, *68*(9–10), 428–436. <https://doi.org/10.1055/s-0043-117604>
- Abu Sabra, M. A., Al Kalaldehy, M., Khalil, M., Abualruz, H., & Hamdan-Mansour, A. M. (2023). The efficacy of using psychotherapy treatments for obsessive–compulsive disorder on minimizing suicidal thoughts and behaviours: A scoping review. *Clinical Psychology & Psychotherapy*, *30*(5), 913–1189. <https://doi.org/10.1002/cpp.2871>
- Allen, J. J., Parker, A., & Ogles, B. M. (2023). A review of the facilitative interpersonal skills performance task and rating method. *Clinical Psychology: Science and Practice*, *31*(2), 211–223. <https://doi.org/10.1037/cps0000187>
- Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of clinical psychology*, *65*(7), 755–768. <https://doi.org/10.1002/jclp.20583>
- Anderson, T., & Patterson, C. (2013). *Facilitative interpersonal skill task and rating method*. Ohio University.
- Anderson, T., Crowley, M. E. J., Himawan, L., Holmberg, J. K., & Uhlin, B. D. (2015). Therapist facilitative interpersonal skills and training status: A randomized clinical trial on alliance and outcome. *Psychotherapy Research*, *26*(5), 511–529. <https://doi.org/10.1080/10503307.2015.1049671>
- Anderson, T., McClintock, A. S., Himawan, L., Song, X., & Patterson, C. L. (2016). A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *Journal of Consulting and Clinical Psychology*, *84*(1), 57–66. <https://doi.org/10.1037/ccp0000060>
- Anderson, T., Perlman, M. R., McCarrick, S. M., & McClintock, A. S. (2019). Modeling therapist responses with structured practice enhances facilitative interpersonal skills. *Journal of Clinical Psychology*, *76*(4), 659– 675. <https://doi.org/10.1002/jclp.22911>
- Anderson, T., & Perlman, M. R. (2022). Therapist and client facilitative interpersonal skills in psychotherapy. In J. N. Fuertes (Ed.), *The other side of psychotherapy: Understanding clients' experiences and contributions in treatment* (pp. 99–124). American Psychological Association. <https://doi.org/10.1037/0000303-005>

- Anderson, T., Stone, S. J., Angus, L., & Weibel, D. T. (2022). Double trouble: Therapists with low facilitative interpersonal skills and without training have low in-session experiential 49 processes. *Psychotherapy Research, 32*(1), 65-77.
<https://doi.org/10.1080/10503307.2021.1913293>
- Axelsson, S., Kihlberg, S., Davis, P., & Nyström, M. B. (2023). Psychotherapy students' experiences of supervisee-centred supervision based on deliberate practice, feedback-informed treatment and self-compassion. *Counselling and Psychotherapy Research, 24*(2), 719–733. <https://doi.org/10.1002/capr.12708>
- Bachelor, A. (2013). Clients' and therapists' views of the therapeutic alliance: Similarities, differences and relationship to therapy outcome. *Clinical psychology & psychotherapy, 20*(2), 118-135. <https://doi.org/10.1002/cpp.792>
- Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance–outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75*(6), 842–852.
<https://doi.org/10.1037/0022-006X.75.6.842>
- Baldwin, S. A., & Imel, Z. E. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 258–297). John Wiley & Sons.
- Banner, S. E., Rock, A. J., Cosh, S. M., Schutte, N., & Rice, K. (2025). Self-reflection on competence: Metacognitive process and barriers of self-assessment in psychologists. *Advances in Health Sciences Education, 30*, 1–21. <https://doi.org/10.1007/s10459-025-10434-7>
- Barkowski, S., Schwartze, D., Strauss, B., Burlingame, G. M., & Rosendahl, J. (2020). Efficacy of group psychotherapy for anxiety disorders: A systematic review and meta-analysis. *Psychotherapy Research, 30*(8), 965–982.
<https://doi.org/10.1080/10503307.2020.1729440>
- Barkham, M., Lutz, W., Lambert, M. J., & Saxon, D. (2017). Therapist effects, effective therapists, and the law of variability. In L. G. Castonguay & C. E. Hill (Eds.), *How and why are some therapists better than others?: Understanding therapist effects* (pp. 13– 36). American Psychological Association. <https://doi.org/10.1037/0000034-002>
- 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* (7th ed., pp. 135–189). Wiley.

- Barkham, M., de Jong, K., 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>
- Barrett-Naylor, R., Malins, S., Levene, J., Biswas, S., Mays, C., & Main, G. (2020). Brief training in psychological assessment and interventions skills for cancer care staff: a mixed methods evaluation of deliberate practice techniques. *Psycho-Oncology, 29*(11), 1786-1793. <https://doi.org/10.1002/pon.5393>
- Bennett-Levy, J. (2019). Why therapists should walk the talk: The theoretical and empirical case for personal practice in therapist training and professional development. *Journal of Behavior Therapy and Experimental Psychiatry, 62*, 133–145. <https://doi.org/10.1016/j.jbtep.2018.08.004>
- Bergin, A. E. (1963). The effects of psychotherapy: Negative results revisited. *Journal of Counseling Psychology, 10*(3), 244–250.
- Bohart, A. C., & Tallman, K. (2010). Clients: The neglected common factor in psychotherapy. In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 83–111). American Psychological Association.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*(3), 252–260. <https://doi.org/10.1037/h0085885>
- Caçador, S., Sousa, D., & Cooper, M. (2023). Inside the consulting room of a highly effective therapist: An analysis of first sessions. *Counselling and Psychotherapy Research, 23*(4), 1–11. <https://doi.org/10.1002/capr.12705>
- Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology, 19*(2), 151–165. <https://doi.org/10.1002/acp.1106>
- Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy, 52*(3), 337–345. <https://doi.org/10.1037/pst0000015>
- Clements-Hickman, A. L., & Reese, R. J. (2020). Improving therapists' effectiveness: Can deliberate practice help? *Professional Psychology: Research and Practice, 51*(6), 606–612. <https://doi.org/10.1037/pro0000318>
- Clements-Hickman, A. L., & Harris, K. R. (2024). Deliberate practice for psychotherapy skills: Recommendations and implications based on the state of the science.

- Professional Psychology: Research and Practice*, 55(6), 563–572.
<https://doi.org/10.1037/pro0000592>
- Cuijpers, P., Van Straten, A., Hollon, S. D., Bockting, C., Berking, M., & Andersson, G. (2013). Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis. *BMJ Open*, 3(4). <https://doi.org/10.1136/bmjopen-2012-002542>
- Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The role of common factors in psychotherapy outcomes. *Annual Review of Clinical Psychology*, 15, 207–231. <https://doi.org/10.1146/annurev-clinpsy-050718-095424>
- Cuijpers, P., Noma, H., Karyotaki, E., Vinkers, C. H., Cipriani, A., & Furukawa, T. A. (2020). A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. *World Psychiatry*, 19(1), 92–107. <https://doi.org/10.1002/wps.20701>
- Cuijpers, P., Quero, S., Noma, H., Ciharova, M., Miguel, C., Karyotaki, E., Cipriani, A., Cristea, I.A., Furukawa, T. (2021). Psychotherapies for depression: a network metaanalysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry*, 20(2), 283–293. <https://doi.org/10.1002/wps.20860>
- de Jong, K., Conijn, J. M., Gallagher, R. A., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, 85, 1-19. <https://doi.org/10.1016/j.cpr.2021.102002>
- Delgadillo, J., Deisenhofer, A.-K., Probst, T., Shimokawa, K., Lambert, M. J., & Kleinstäuber, M. (2022). Progress feedback narrows the gap between more and less effective therapists: A therapist effects meta-analysis of clinical trials. *Journal of Consulting and Clinical Psychology*, 90(7), 559–567. <https://doi.org/10.1037/ccp0000747>
- Di Bartolomeo, A. A., Shukla, S., Westra, H. A., Shekarak Ghashghaei, N., & Olson, D. A. (2021). Rolling with resistance: A client language analysis of deliberate practice in continuing education for psychotherapists. *Counselling and Psychotherapy Research*, 21(2), 433-441. <https://doi.org/10.1002/capr.12335>
- Dixon, L. J., & Linardon, J. (2020). A systematic review and meta-analysis of dropout rates from dialectical behaviour therapy in randomized controlled trials. *Cognitive Behaviour Therapy*, 49(3), 181-196. <https://doi.org/10.1080/16506073.2019.1620324>

- Edmondstone, C., Pascual-Leone, A., Soucie, K., & Kramer, U. (2023). Therapist effects on outcome: Meaningful differences exist early in training. *Training and Education in Professional Psychology, 17*(2), 149–157. <https://doi.org/10.1037/tep0000402>
- Eielsen, M., Ulvenes, P. G., Røssberg, J. I., Høstmælingen, A., Soma, C. S., & Wampold, B. E. (2022). The Effectiveness of an Intensive Inpatient Psychotherapy Program for Chronic Depression: A naturalistic comparison with wait list. *BMC Psychiatry, 22*. <https://doi.org/10.1186/s12888-022-04381-5>
- Elliott, R., Bohart, A. C., Watson, J. C. & Greenberg, L. S. (2011). Empathy. *Psychotherapy: Theory, Research, Practice & Training, 48*(1), 43–49. <https://doi.org/10.1037/a0022187>
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406. <https://doi.org/10.1037/0033-295X.100.3.363>
- Ericsson, A., & Pool, R. (2016). *Peak: Secrets from the new science of expertise*. Houghton Mifflin Harcourt.
- Ericsson, K. A., Hoffman, R. R., Kozbelt, A., & Williams, A. M. (Eds.). (2018). *The Cambridge handbook of expertise and expert performance* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/9781316480748>
- Ericsson, K. A., & Harwell, K. W. (2019). Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future research. *Frontiers in Psychology, 10*, 2396. <https://doi.org/10.3389/fpsyg.2019.02396>
- Ericsson, K. A. (2020). Towards a science of the acquisition of expert performance in sports: Clarifying the differences between deliberate practice and other types of practice. *Journal of sports sciences, 38*(2), 159-176. <https://doi.org/10.1080/02640414.2019.1688618>
- Ericsson, K. A. (2021). Given that the detailed original criteria for deliberate practice have not changed, could the understanding of this complex concept have improved over time? A response to Macnamara and Hambrick (2020). *Psychological Research, 85*(3), 1114- 1120. <https://doi.org/10.1007/s00426-020-01368-3>
- Eysenck, H. J. (1952). The effects of psychotherapy: an evaluation. *Journal of Consulting Psychology, 16*(5), 319–324. <https://doi.org/10.1037/h0063633>
- Farber, B. A., Suzuki, J. Y., & Ort, D. (2022). Positive regard and psychotherapy: Controversies, criticisms, and conclusions. In *Understanding and enhancing positive*

- regard in psychotherapy: *Carl Rogers and beyond*. (pp. 203–223). American Psychological Association. <https://doi.org/10.1037/0000312-010>
- Fernandez, E., Salem, D., Swift, J. K., & Ramtahal, N. (2015). Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. *Journal of Consulting and Clinical Psychology, 83*(6), 1108–1122. <https://doi.org/10.1037/ccp0000044>
- Fernandez, D., Vigo, D., Sampson, N. A., Hwang, I., Aguilar-Gaxiola, S., Al-Hamzawi, A. O., Alonso, J., Andrade, L. H., Bromet, E. J., de Giromalo, G., de Jonge, P., Florescu, S., Gureje, O., Hinkov, H., Hu, C., Karam, E. G., Karam, G., Kawakami, N., Kiejna, A., ... Haro, J. M. (2021). Patterns of care and dropout rates from outpatient mental healthcare in low-, middle-and high-income countries from the World Health Organization's World Mental Health Survey Initiative. *Psychological medicine, 51*(12), 2104-2116.. <https://doi.org/10.1017/S0033291720000884>
- Ferreira, L., Guerra, C., & Vieira-Coelho, M. A. (2023). Predictors of psychotherapy dropout in patients with borderline personality disorder: A systematic review. *Clinical Psychology & Psychotherapy, 1*-8. <https://doi.org/10.1002/cpp.2888>
- Frank, J. D., & Frank, J. B. (1991). *Persuasion and healing: A comprehensive study of psychotherapy* (3rd ed.). Johns Hopkins University Press.
- Gelso, C. (2014). A tripartite model of the therapeutic relationship: Theory, research, and practice. *Psychotherapy research, 24*(2), 117-131. <https://doi.org/10.1080/10503307.2013.845920>
- Gjerstad, S. F., Nordin, L., Poulsen, S., Spadaro, E. F. A., & Palic, S. (2024). How is trauma-focused therapy experienced by adults with PTSD? A systematic review of qualitative studies. *BMC Psychology, 12*(1), 1–19. <https://doi.org/10.1186/s40359-024-01588-x>
- Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T., & Wampold, B. E. (2016). Do psychotherapists improve with time and experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counseling Psychology, 63*(1), 1–11. <https://doi.org/10.1037/cou0000131>
- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of clinical psychology, 61*(2), 155-163. <https://doi.org/10.1002/jclp.20108>

- Hartmann, A., Joos, A., Orlinsky, D. E., & Zeeck, A. (2014). Accuracy of therapist perceptions of patients' alliance: Exploring the divergence. *Psychotherapy Research, 25*(4), 408–419. <https://doi.org/10.1080/10503307.2014.927601>
- Hashimoto, D. A., Sirimanna, P., Gomez, E. D., Beyer-Berjot, L., Ericsson, K. A., Williams, N. N., Darzi, A., & Aggarwal, R. (2015). Deliberate practice enhances quality of laparoscopic surgical performance in a randomized controlled trial: From arrested development to expert performance. *Surgical Endoscopy, 29*(11), 3154–3162. <https://doi.org/10.1007/s00464-014-4042-4>
- Hatfield, D., McCullough, L., Frantz, S. H. B., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology and Psychotherapy, 17*(1), 25–32. <https://doi.org/10.1002/cpp.656>
- Heinonen, E., & Nissen-Lie, H. A. (2020). The professional and personal characteristics of effective psychotherapists: A systematic review. *Psychotherapy Research, 30*(4), 417–432. <https://doi.org/10.1080/10503307.2019.1620366>
- Hilbert, A., Petroff, D., Herpertz, S., Pietrowsky, R., Tuschen-Caffier, B., Vocks, S., & Schmidt, R. (2019). Meta-analysis of the efficacy of psychological and medical 54 treatments for binge-eating disorder. *Journal of Consulting and Clinical Psychology, 87*(1), 91–105. <https://doi.org/10.1037/ccp0000358>
- Hill, C. E., Kivlighan, D. M. III, Rousmaniere, T., Kivlighan, D. M., Jr., Gerstenblith, J. A., & Hillman, J. W. (2020). Deliberate practice for the skill of immediacy: A multiple case study of doctoral student therapists and clients. *Psychotherapy, 57*(4), 587–597. <https://doi.org/10.1037/pst0000247>
- Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38*(2), 139–149. <https://doi.org/10.1037/0022-0167.38.2.139>
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy, 48*(1), 9–16. <https://doi.org/10.1037/a0022186>
- Høstmælingen, A., Ulvenes, P., Nissen-Lie, H. A., Eielsen, M., & Wampold, B. E. (2021). Comparing outcomes in chronic depression following inpatient psychotherapy for patients continuing versus discontinuing antidepressant medication. *Clinical Psychology & Psychotherapy, 28*(5), 1111–1127. <https://doi.org/10.1002/cpp.2561>
- Hunt, E. A., Duval-Arnould, J. M., Nelson-McMillan, K. L., Bradshaw, J. H., Diener-West, M., Perretta, J. S., & Shilkofski, N. A. (2014). Pediatric resident resuscitation skills

- improve after “Rapid Cycle Deliberate Practice” training. *Resuscitation*, 85(7), 945–951. <https://doi.org/10.1016/j.resuscitation.2014.02.025>
- Iliakis, E. A., Ilagan, G. S., & Choi-Kain, L. W. (2021). Dropout rates from psychotherapy trials for borderline personality disorder: A meta-analysis. *Personality Disorders: Theory, Research, and Treatment*, 12(3), 193–206. <https://doi.org/10.1037/per0000453>
- Imel, Z. E., Malterer, M. B., McKay, K. M., & Wampold, B. E. (2008). A meta-analysis of psychotherapy and medication in unipolar depression and dysthymia. *Journal of affective disorders*, 110(3), 197–206. <https://doi.org/10.1016/j.jad.2008.03.018>
- Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review*, 67, 78–93. <https://doi.org/10.1016/j.cpr.2018.08.004>
- Kappelman, N., Rein, M., Fietz, J., Mayberg, H. S., Craighead, W. E., Dunlop, B. W., Nemeroff, C. B., Keller, M., Klein, D. N., Arnow, B. A., Husain, N., Jarrett, R. B., Vittengl, J. R., Menchetti, M., Parker, G., Barber, J. P., Bastos, A. G., Dekker, J., Peen, J., & Keck, M. E. (2020). Psychotherapy or medication for depression? Using individual symptom meta-analyses to derive a Symptom-Oriented Therapy (SOt) metric for a personalised psychiatry. *BMC Medicine*, 18(1), 1–18. <https://doi.org/10.1186/s12916-020-01623-9>
- Klatte, R., Strauss, B., Flückiger, C., Färber, F., & Rosendahl, J. (2023). Defining and assessing adverse events and harmful effects in psychotherapy study protocols: A systematic review. *Psychotherapy*, 60(1), 130–148. <https://doi.org/10.1037/pst0000359>
- Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, 125(4), 331–359. <https://doi.org/10.1037/0096-3445.125.4.331>
- Lagerberg, H., Boswell, J. F., Constantino, M. J., Andersson, G., & Carlbring, P. (2024). Does practice make perfect? The effects of an eight-week manualized deliberate practice course with peer feedback on patient-rated working alliance in adults: A pilot randomized controlled trial. *Clinical Psychology in Europe*, 6(3), 1–19. <https://doi.org/10.32872/cpe.12353>
- Lambert, M. J., Harmon, C., Slade, K., Whipple, J. L., & Hawkins, E. J. (2005). Providing feedback to psychotherapists on their patients' progress: clinical results and practice

- suggestions. *Journal of clinical psychology*, 61(2), 165-174.
<https://doi.org/10.1002/jclp.20113>
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (2nd ed., pp. 203–223). Oxford University Press.
- Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, 50(1), 42–51. <https://doi.org/10.1037/a0030682>
- Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy*, 55(4), 520–537. <https://doi.org/10.1037/pst0000167>
- Larsson, J., Werthén, D., Carlsson, J., Salim, O., Davidsson, E., Vaz, A., Sousa, D., & Norberg, J. (2023). Does deliberate practice surpass didactic training in learning empathy skills? – A randomized controlled study. *Nordic Psychology*, 77(1), 39–52. <https://doi.org/10.1080/19012276.2023.2247572>
- Laska, K. M., Gurman, A. S., & Wampold, B. E. (2013). Expanding the lens of evidence-based practice in psychotherapy: A common factors perspective. *Psychotherapy*, 51(4), 467–481. <https://doi.org/10.1037/a0034332>
- Leichsenring, F., Steinert, C., Rabung, S., & Ioannidis, J. P. A. (2022). The efficacy of psychotherapies and pharmacotherapies for mental disorders in adults: An umbrella review and meta-analytic evaluation of recent meta-analyses. *World Psychiatry*, 21(1), 133–145. <https://doi.org/10.1002/wps.20941>
- Linardon, J., Fitzsimmons-Craft, E. E., Brennan, L., Barillaro, M., & Wilfley, D. E. (2019). Dropout from interpersonal psychotherapy for mental health disorders: A systematic review and meta-analysis. *Psychotherapy Research*, 29(7), 870-881. <https://doi.org/10.1080/10503307.2018.1497215>
- Longley, M., Kästner, D., Daubmann, A., Hirschmeier, C., Strauß, B., & Gumz, A. (2023). Prospective psychotherapists' bias and accuracy in assessing their own facilitative interpersonal skills. *Psychotherapy*, 60(4), 1–11. <https://doi.org/10.1037/pst0000506>
- Macdonald, J., & Mellor-Clark, J. (2015). Helpers' self-assessment biases before and after helping skills training. *The European Journal of Counselling Psychology*, 3(1), 1–17. <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01377/full>

- Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science, 25*(8), 1608-1618. <https://doi.org/10.1177/0956797614535810>
- Macnamara, B. N., Moreau, D., & Hambrick, D. Z. (2016). The relationship between deliberate practice and performance in sports: A meta-analysis. *Perspectives on Psychological Science, 11*(3), 333–350. <https://doi.org/10.1177/17456916166635591>
- Mahon, D. (2022). A scoping review of deliberate practice in the acquisition of therapeutic skills and practices. *Counselling and Psychotherapy Research, 23*(4), 965- 981. <https://doi.org/10.1002/capr.12601>
- Mahon, D. (2024). “Don't just magically assume you're going to get better as a therapist as time goes on”: An exploration of the use of deliberate practice. *Counselling and Psychotherapy Research, 24*(3), 913–924. <https://doi.org/10.1002/capr.12746>
- McGaghie, W. C., Issenberg, B., Cohen, E. R., Barsuk, J. H., & Wayne, D. B. (2011). Does Simulation-Based Medical Education With Deliberate Practice Yield Better Results Than Traditional Clinical Education? A Meta-Analytic Comparative Review of the Evidence. *Academic Medicine, 86*(6), 706–711. <https://doi.org/10.1097/ACM.0b013e318217e119>
- Miller, S. D., Hubble, M. A., and Duncan, B. L. (2007). Supershrinks: Learning from the field’s most effective practitioners. *Psychotherapy Networker, 31*, 26–35, 56.
- Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2013). The outcome of psychotherapy: Yesterday, today, and tomorrow. *Psychotherapy, 50*(1), 88–97. <https://doi.org/10.1037/a0031097>
- Miller, S. D., Chow, D., Wampold, B. E., Hubble, M. A., Del Re, A. C., Maeschalck, C., & Bargmann, S. (2020). To be or not to be (an expert)? Revisiting the role of deliberate practice in improving performance. *High Ability Studies, 31*(1), 5–15. <https://doi.org/10.1080/13598139.2018.1519410>
- Moritz, S., Nestoriuc, Y., Rief, W., Klein, J. P., Jelinek, L., & Peth, J. (2019). It can’t hurt, right? Adverse effects of psychotherapy in patients with depression. *European Archives of Psychiatry and Clinical Neuroscience, 269*(5), 577–586. <https://doi.org/10.1007/s00406-018-0931-1>
- Munder, T., Flückiger, C., Leichsenring, F., Abbass, A. A., Hilsenroth, M. J., Luyten, P., Rabung, S., Steinert, C., & Wampold, B. E. (2019). Is psychotherapy effective? A reanalysis of treatments for depression. *Epidemiology and Psychiatric Sciences, 28*(3), 268-274. <https://doi.org/10.1017/S2045796018000355>

- Newman, D. S., Villarreal, J. N., Gerrard, M. K., McIntire, H., Barrett, C. A., & Kaiser, L. T. (2022). Deliberate practice of consultation communication skills: A randomized controlled trial. *School Psychology, 37*(3), 225- 235.
<https://doi.org/10.1037/spq0000494>
- Nissen-Lie, H. A., Rønnestad, M. H., Høglend, P. A., Havik, O. E., Solbakken, O. A., Stiles, T. C., & Monsen, J. T. (2015). Love yourself as a person, doubt yourself as a therapist? *Clinical Psychology & Psychotherapy, 24*(1), 48–60.
<https://doi.org/10.1002/cpp.1977>
- Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy, 55*(4), 303–315. <https://doi.org/10.1037/pst0000193>
- Olsson, M., Gustafsson, M., Carlsson, J., Rousmaniere, T., Bergbom, S., & Norberg, J. (2024). A pilot study of a deliberate practice intervention for therapist trainees. *Nordic Psychology, 77*(2), 165–184. <https://doi.org/10.1080/19012276.2024.2313532>
- Owen, J., & Hilsenroth, M. J. (2014). Treatment adherence: The importance of therapist flexibility in relation to therapy outcomes. *Journal of Counseling Psychology, 61*(2), 280–288. <https://doi.org/10.1037/a0035753>
- Perlman, M. R., Anderson, T., Finkelstein, J. D., Foley, V. K., Mimnaugh, S., Gooch, C. V., David, K. C, Martin, S. J., & Safran, J. D. (2023). Facilitative interpersonal relationship training enhances novices’ therapeutic skills. *Counselling Psychology Quarterly, 36*(1), 25-40. <https://doi.org/10.1080/09515070.2022.2049703>
- Perlman, M. R., Anderson, T., Foley, V. K., Mimnaugh, S., & Safran, J. D. (2020). The impact of alliance-focused and facilitative interpersonal relationship training on therapist skills: An RCT of brief training. *Psychotherapy Research, 30*(7), 871-884.
<https://doi.org/10.1080/10503307.2020.1722862>
- Pereira, G. L., Trujillo-Sánchez, C., Alonso-Vega, J., Echevarría-Escalante, D., & FroxánParga, M. X. (2023). What do we know about highly effective therapists? *A systematic review. Annals of psychology, 39*(1), 10-19.
<https://doi.org/10.6018/analesps.499371>
- Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2014). The influence of deliberate practice on musical achievement: a meta-analysis. *Frontiers in Psychology, 5*, 646.
<https://doi.org/10.3389/fpsyg.2014.00646>
- Rahmani, F., Abbass, A., Hemmati, A., Mirghaed, S. R., & Ghaffari, N. (2020). The efficacy of intensive short-term dynamic psychotherapy for social anxiety disorder: randomized trial and substudy of emphasizing feeling versus defense work. *The*

- Journal of nervous and mental disease*, 208(3), 245-251.
<https://doi.org/10.1097/NMD.0000000000001097>
- Ricks, D. F. (1974). Supershrink: Methods of a therapist judged successful on the basis of adult outcomes of adolescent patients. In D. F. Ricks, A. Thomas, & M. Roff (Eds.), *Life history research in psychopathology: III*. University of Minnesota Press
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, 21(2), 95–103.
<https://doi.org/10.1037/h0045357>
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry*, 6(3), 412–415.
<https://doi.org/10.1037/h0061172>
- Rosenzweig, S. (1954). A transvaluation of psychotherapy: a reply to Hans Eysenck. *The Journal of Abnormal and Social Psychology*, 49(2), 298–304.
<https://doi.org/10.1037/h0061172>
- Rousmaniere, T. (2017). *Deliberate practice for psychotherapists: A guide to improving clinical effectiveness*. Routledge.
- Rousmaniere, T., Goodyear, R. K., Miller, S. D., & Wampold, B. E. (Eds.). (2017). *The cycle of excellence: Using deliberate practice to improve supervision and training*. John Wiley & Sons.
- Rousmaniere, T. (2019). *Mastering the inner skills of psychotherapy: A deliberate practice manual*. Gold Lantern Press.
- Safran, J. D., & Muran, J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. Guilford Press.
- Safran, J. D., Muran, J. C., Samstag, L. W., & Winston, A. (2005). Evaluating alliance-focused intervention for potential treatment failures: A feasibility study and descriptive analysis. *Psychotherapy: Theory, Research, Practice, Training*, 42(4), 512–531. <https://doi.org/10.1037/0033-3204.42.4.512>
- Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. In *Psychotherapy relationships that work: Evidence-based responsiveness*, 2nd ed. (pp. 224–238). <https://doi.org/10.1093/acprof:oso/9780199737208.003.0011>
- Safran, J. D., & Kraus, J. (2014). Alliance ruptures, impasses, and enactments: A relational perspective. *Psychotherapy*, 51(3), 381–387. <https://doi.org/10.1037/a0036815>
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a

- psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78(3), 298. <https://doi.org/10.1037/a0019247>
- Shukla, S., Di Bartolomeo, A. A., Westra, H. A., Olson, D. A., & Shekarak Ghashghaei, N. (2021). The impact of a deliberate practice workshop on therapist demand and support behavior with community volunteers and simulators. *Psychotherapy*, 58(2), 186–195. <https://doi.org/10.1037/pst0000333>
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(9), 752–760. <https://doi.org/10.1037/0003-066X.32.9.752>
- Strauss, B., Gawlytta, R., Schleu, A., & Frenzl, D. (2021). Negative effects of psychotherapy: Estimating the prevalence in a random national sample. *BJPsych Open*, 7(6), 1-8. <https://doi.org/10.1192/bjo.2021.1025>
- Spengler, P. M., & Pilipis, L. A. (2015). A comprehensive meta-reanalysis of the robustness of the experience-accuracy effect in clinical judgment. *Journal of Counseling Psychology*, 62(3), 360–378. <https://doi.org/10.1037/cou0000065>
- Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 80(4), 547. <https://doi.org/10.1037/a0028226>
- Swift, J. K., & Greenberg, R. P. (2014). A Treatment by disorder meta-analysis of dropout from Psychotherapy. *Journal of Psychotherapy Integration*, 24(3), 193–207. <https://doi.org/10.1037/a0037512>
- Sousa, D. (2017). *Investigação científica em psicoterapia e prática psicoterapêutica: Os dados da investigação mais relevantes para os clínicos*. Fim de Século
- Strupp, H. H. (1963). The outcome problem in psychotherapy revisited. *Psychotherapy: Theory, Research & Practice*, 1(1), 1–13. <https://doi.org/10.1037/h0088565>
- Timulak, L. (2010). Significant events in psychotherapy: An update of research findings. *Psychology & Psychotherapy: Theory, Research & Practice*, 83(4), 421–447. <https://doi.org/10.1348/147608310X499404>
- Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist*, 69(3), 218–229. <https://doi.org/10.1037/a0035099>
- Truax, C. B., & Carkhuff, R. R. (1967). *Toward effective counseling and psychotherapy: Training and practice*. Aldine.

- Tryon, G. S., Blackwell, S. C., & Hammel, E. F. (2007). A meta-analytic examination of client– therapist perspectives of the working alliance. *Psychotherapy Research, 17*(6), 629-642. <https://doi.org/10.1080/10503300701320611>
- Tryon, G. S., & Winograd, G. (2011). Goal consensus and collaboration. *Psychotherapy, 48*(1), 50–57. <https://doi.org/10.1037/a0022061>
- van Dijk, D. A., Deen, M. L., van den Boogaard, T. M., Ruhé, H. G., Spijker, J., & Peeters, F. P. M. L. (2023). Prevalence and prediction of dropout during depression treatment in routine outpatient care: an observational study. *European archives of psychiatry and clinical neuroscience, 273*(5), 1151-1161. <https://doi.org/10.1007/s00406-022-01499-1>
- Vaz, A., & Rousmaniere, T. (2022). Clarifying deliberate practice for mental health training. *Practice Innovations, 7*(2), 143–156.
- Vaz, A., Sousa, D., & Afonseca, M. (2023). *Treino em psicoterapia. Como usar a prática deliberada para melhorar resultados clínicos*. Pactor.
- Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of selfassessment bias in mental health providers. *Psychological reports, 110*(2), 639-644. <https://doi.org/10.2466/02.07.17.PR0.110.2.639-6>
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Lawrence Erlbaum Associates.
- Wampold, B. E., & Brown, G. (2005). Estimating therapist variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*(5), 914-923. <https://doi.org/10.1037/0022-006X.73.5.914>
- Wampold, B. E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist, 62*(8), 857–873. <https://doi.org/10.1037/0003-066X.62.8.857>
- Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World psychiatry, 14*(3), 270-277. <https://doi.org/10.1002/wps.20238>
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.
- Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A., & Burlingame, G. M. (2010). Youth psychotherapy change trajectories and outcomes in usual care: Community mental health versus managed care settings. *Journal of Consulting and Clinical Psychology, 78*(2), 144. <https://doi.org/10.1037/a0018544>

- Weber, M., Schumacher, S., Hannig, W., Barth, J., Lotzin, A., Schäfer, I., Ehring, T., & Kleim, B. (2021). Long-term outcomes of psychological treatment for posttraumatic stress disorder: a systematic review and meta-analysis. *Psychological Medicine*, *51*(9), 1420–1430. <https://doi.org/10.1017/S003329172100163X>
- Westra, H. A., Norouzian, N., Poulin, L., Coyne, A., Constantino, M. J., Hara, K., Olson, D., & Antony, M. M. (2021). Testing a deliberate practice workshop for developing appropriate responsivity to resistance markers. *Psychotherapy*, *58*(2), 175– 185. <https://doi.org/10.1037/pst0000311>
- Yamin, J. B., Cannoy, C. N., Gibbins, K. M., Krohner, S., Rapport, L. J., Trentacosta, C. J., Zeman, L. L., & Lumley, M. A. (2023). Experiential training of mental health graduate students in emotional processing skills: A randomized, controlled trial. *Psychotherapy*, *60*(4), 512–524. <https://doi.org/10.1037/pst0000495>
- Young, J., & Maack, D. J. (2021). Incorporating deliberate practice in the process of continual therapeutic skill development. *Cognitive and Behavioral Practice*, *28*(2), 267–280. <https://doi.org/10.1016/j.cbpra.2020.09.009>

Anexo II

Consentimento Informado



Consentimento Informado

A Especialização Avançada em Psicoterapia (EAP) vai incentivar e promover a articulação entre a formação, a prática clínica e a investigação científica. Vimos solicitar a sua colaboração para um primeiro estudo. Esta investigação pretende estudar a possibilidade de a formação em prática deliberada em psicoterapia melhorar as competências interpessoais dos psicoterapeutas. A sua participação consistirá em: a) responder a 9 vídeos que correspondem a cenários fictícios de sessões de psicoterapia, na plataforma *Skillssetter*; b) uma segunda resposta aos mesmos vídeos depois do seminário de prática deliberada; c) a realização de uma entrevista qualitativa após o seminário de prática deliberada.

Em virtude desta colaboração clarifica-se os seguintes aspetos:

1. O estudo destina-se exclusivamente a fins de investigação e publicação científica.
2. A sua participação é totalmente anónima e voluntária, podendo desistir de participar a qualquer momento e solicitar a eliminação das gravações.
3. A cotação dos vídeos será realizada por três investigadores. Todos os investigadores estão obrigados ao mesmo dever de sigilo e ao cumprimento de um código deontológico.
4. Não são previsíveis efeitos adversos por participar neste estudo. Ao fazê-lo contribui para a investigação, no âmbito da prática deliberada.

Para qualquer questão ou dúvida sobre o estudo poderá sempre contactar a Daniel Sousa (daniel.sousa@ispa.pt) da Direcção da EAP.

Eu _____ compreendo as condições apresentadas e concordo com os termos.

Anexo III

Questionário Sociodemográfico

Questionário Sociodemográfico	
Especialização Avançada em Psicoterapia	
1. Qual é a sua idade? *	Texto de resposta curta
2. Com que género se identifica? *	<input type="radio"/> Feminino <input type="radio"/> Masculino <input type="radio"/> Prefiro não dizer <input type="radio"/> Outra:
3. Nº de anos de prática clínica: *	Texto de resposta curta
4. Nº de horas semanais de prática clínica: *	Texto de resposta curta

4. Nº de horas semanais de prática clínica: *

Texto de resposta curta

5. Qual o modelo teórico que mais utiliza? *

Psicodinâmico

Humanista-Existencial

Sistémico-Familiar

Cognitivo-Comportamental

Integrativo

Outra:

Obrigado

Descrição (opcional)

Título da imagem



Anexo IV

Facilitative Interpersonal Skills Rating Scale

Facilitative Interpersonal Skills Rating Scale
Timothy Anderson & Candace Patterson

Ratings for each item are made on a 5 point Likert-type scale. The rating scale ranges from Not Characteristic ("1") to Extremely Characteristic ("5"). More elaborate descriptions of each rating level are provided to help identify the correct rating level.

Response Set: Clearly, people differ in how they evaluate the skillfulness of helper interventions. The descriptions of these items are lengthy in order to provide the context for what is intended by each particular FIS domain / item. Items are written in order to assist the rater in using specific evidence from the tape to inform rating decisions. In rating a response, instances of an "average" level of a helping behavior would merit a rating of 3. Ratings of 3 are thought of as the default rating for all items and are considered ordinary helping or facilitative interpersonal skills. Thus, a 3 should be the starting point for ratings and the participant's response may influence you to increase or decrease your rating from a 3.

PART I

1. Verbal Fluency. This item is a rating of the extent to which the participant is verbally capable and at-ease in communicating. The response is delivered in a relaxed manner and without significant signs of anxiety (e.g., broken speech, extended and awkward pauses, and clarity in communication). However, the content of what is said is not rated, but rather how it is spoken.

- 5 The participant is at great ease and communicates ideas with no anxiety, reflecting a desire to "approach" the other. The verbal quality of the response may have a "melodic," rhythmical quality and is easy to follow; the response is fluent.
- 4 The response is fluent, and there is little that is difficult to follow.
- 3 A moderate level of verbal fluency indicates that the participant's response is conversational and mostly easy to follow.
- 2 Fluency is disrupted by the participant's anxiety and avoidance of verbal expression. The respondent may be obviously anxious and struggling to formulate a response. At times, the communication may be choppy, even halting. [Note: In some rare instances a response could represent an avoidance of the interpersonal situation through anxious rambling. It would need to be clear that the participant's rambles are the result of anxiety over communicating with another].
- 1 The participant has great difficulty verbalizing a response (e.g., obviously anxious, sounds shaky or timid), reflecting a clear avoidance or anxiety. The participant may lack confidence and is clearly uncertain or even difficult to follow.

2. Hope & Positive Expectations. This item rates expressions of hope, optimism, and positive expectations for change. Staats (1989, 2001) defines hope as the interaction between wishes and expectations. The interpersonal skills needed for hope involve facilitating a) personal *agency* and b) building the *pathways* needed for attaining desired goals and expectations (Steed, 2002). Hope is related to persuasiveness and collaboration in the sense that hope and positive expectations are often built through offering a rationale, friendliness, and enthusiasm. As defined here, hope focuses more on building client agency for actions that will facilitate meeting the client's goals whereas persuasion is based more on a plausible explanation (which may or may not include hope).

- 5 The participant's response expresses clear hope about the client's future and/or positive expectations about therapeutic work. In addition, for a response to be coded as a "5" there needs to be an allusion to building the client's *agency* as well as *how* the client might participate or do something that will help move toward his/her desired goals (i.e., pathways).
- 4 A general sense of optimism about the client's situation is detected. Specifically, the participant's response is directed toward building the client's *agency* OR facilitating the building of pathways to meet the client's goals.
- 3 The response is ordinary OR the optimism of the response is not discernable. There may be some *hopefulness* expressed, but with little confidence or reason for being hopeful.
- 2 The participant responds with some hopelessness, including subtle expressions of feeling unable to help the client.
- 1 The participant's response is hopeless or is even pessimistic. For example, the participant may address only issues or concerns beyond the control of the other or subtly suggests that the other cannot change or improve his/her problems.

3. Persuasiveness. Persuasiveness is the capacity to induce the other to accept a view that may be different from his or her own view. It involves that ability to convey a clear, organized understanding about the meaning of the other's source of distress. Persuasiveness implies an ability to communicate what Jerome Frank called a "believable myth." This capacity implies that the persuasive therapist must be convincing in communicating this belief-system.

Rating Notes: For rating purposes, the response does not necessarily need to convey an entire world view, but a point of view that is implied to be at least slightly different from the client in the video clip. High ratings require that the participant provide a clear belief in a *point of view* or rationale. It is necessary that the rationale be relevant to the other's problems and at least somewhat novel to the other's experience.

For this item, the rater should disregard personal beliefs about the validity of the participant's rationale, but instead rate the extent to which the participant might persuade another (i.e., ability to "sell" their rationale).

- 5 The participant is highly persuasive. Persuasive persons may speak with great confidence, certainty, and authority. Advice may or may not be given, but the participant must offer some rationale or re-framing of the other's experience.
- 4 The participant speaks persuasively. The rationale may be more implicit and it is even possible that the rationale, though present, may be unclear, superficial, or marginally relevant to the other's problems.
- 3 The participant's response conveys little sense of persuasiveness.
- 2 The participant is unpersuasive. Unpersuasive responses may be characterized by either
 - a) a rationale that lacks *credibility* and there is little reason to believe that it could be convincing. It is important here to try to be aware of your personal biases in judging credibility. As a rule, you can accept most explanations offered as being credible unless there is a clear logical flaw in the *process of explaining* their particular belief.
 - b) a response that is expressed with a lack of confidence, lethargy, or uncertainty by the respondent will be low in persuasiveness. *Even responses that don't contain a rationale may be coded as low in persuasiveness*
- 1 The participant's response is unorganized, incoherent, and difficult to follow. The participant may also not know what to say.

PART II

4. **Emotional Expression.** This item rates the energy and emotion in the participant's response. This item rates the extent to which the participant's response is delivered with effective expressions of emotion.

- 5 There is affect and prosody in the participant's voice. The response is delivered in a highly **emotional and engaging** manner. The primary criterion is that the vocal expression conveys emotion. There may be a more focused delivery of emotional intonations to emphasize meanings that influence other processes (e.g. persuasion). The participant may even be somewhat provocative or challenging in delivering an emotion-based response toward the client in the video clip. However, a "5" should not be rated if the affect is primarily demeaning or hostile toward the other (in which case a "3" would be the maximum rating possible).
- 4 The participant is emotionally expressive at a moderate level. There is more emotion than found in ordinary speech, but it is not as focused in its delivery as the maximum rating of 5.
- 3 The participant has prosody, but it is the amount of emotion that one might find in ordinary conversation.
- 2 The participant may display some sense of interest or curiosity, but the response is **not emotionally engaging**. Prosody is somewhat less than typical to casual conversation.
- 1 The participant speaks with **little or no affect** and may be dull or boring (e.g., speaking in a near monotone voice and without intensity).

5. **Warmth, Acceptance, & Understanding.** This item is a rating of the ability of the participant to care for and accept the other. Therapist behaviors/attitudes that might indicate an absence of acceptance and understanding include: a judgmental attitude, condescension, rudeness, disapproval, guilt-induction, exasperation, or annoyance. Often it will be necessary to avoid rating what the participant is doing (e.g., giving advice), but rate how it is being done. Note that accepting does not necessarily mean approval, but rather a caring attitude and determination to help the other.

- 5 The participant expresses clear and obvious warmth, concern and acceptance. The participant may, for example, make a compassionate attempt to relate to the other's experience.
 - 4 The participant's response is genuinely nonjudgmental and gently explores the other's thoughts, feelings, alternatives for dealing with future situations, etc. The participant appears concerned for and respectful of the client.
 - 3 There is an "ordinary" level of courtesy and warmth in the response OR the participant's opinion of the other may not be clearly discernable from the response.
 - 2 The participant conveys a subtle lack of respect, acceptance, or concern of the other (e.g., sarcasm, exasperation, annoyance).
 - 1 The participant has an obvious lack of respect, acceptance, or warmth for the other (e.g., clearly pejorative comments, judgmental attitude, condescension, disapproval, guilt induction, blaming the other).
-

6. Empathy. The capacity to respond with an expressed understanding of the subjective experience of the client. The response must also convey an *accurate* understanding of the thoughts and emotions expressed in the video clip. Therefore, it is especially important that the rater have an accurate understanding of the client's experiences in the video clips.

- 5 Participant alludes to the client's experience so that it is clear that he/she has not only listened, but obtained an exceptional comprehension of what the other is experiencing. In order to receive a "5" the participant must infer something about the other's experience that is not explicitly stated by the other but is clearly identifiable in the client's nonverbal expression.
- 4 Participant comments accurately on the other's experience but not to the extent required to receive a "5" rating. The distinction between the 4 and 5 ratings are matters of intensity.
- 3 Participant is generally accurate about the other's experience but only perceives the more obvious aspects of the other's experience or concerns.
- 2 Participant does not communicate an awareness or understanding of the other's experience, and/or there are minor distortions of the other's experience. Some aspects of the participant's response may be irrelevant to the other's concerns (when clearly not an attempt to change the other's focus).
- 1 Participant clearly distorts the other's experience. That is, the participant misidentifies a significant component of the other's complaints, beliefs, emotions, etc. Give a rating of 1 if the response indicates a clear disregard of the other's experience.

7. Alliance Bond Capacity. This item rates the participant's capacity to provide a collaborative environment, one in which there is recognition of the need to work with the client jointly on problems.

- 5 Specific actions on the part of the participant help create a collaborative atmosphere. There should be a sense that the participant is attempting to work with the other to create a "we-ness" that is implied in the participant's behavior (e.g., participant checks with the other by asking questions about the "fit" of interpretations, conclusions, goals, etc.).
- 4 Some effort to collaborate is made but not as strong as a "5" (e.g., subtle invitations to engage in working with the client).
- 3 The participant neither undermines nor attempts to enhance a collaborative effort.
- 2 The participant may slightly undermine the building of a collaborative atmosphere, although it may be unintentional or superficial.
- 1 The participant actively undermines a mutual collaboration. The participant may respond in a way that is over-involved or reactive (e.g., moralistic lecturing, "preaching" to the other, assuming all responsibility). The rupture may also involve withdrawal or under-involvement in the participant's response (e.g., putting all the responsibility for change on the other).

8. Alliance Rupture-Repair Responsiveness.

Background: Each client in video clips is expressing an interpersonal issue that involves the patient-therapist relationship. Each video clip places the participant in the middle of alliance rupture episodes. Further, these rupture episodes take place at different locations within the interpersonal circle, which requires interpersonal flexibility for the therapist. The interpersonal problem with the client-therapist relationship threatens the development of the alliance.

This item rates the extent to which the therapist appears responsive to the interpersonal issue. In some cases, the problem is clearly stated as when Suzie angrily berates the therapist for being ineffective. In other cases, the problem is more implicit such as when Lauren idealizes the therapist to the extent of leaving herself overly vulnerable to disappointment.

- 5 Participant makes attempts repair the interpersonal issue by engaging the client in a direct discussion of the immediate moment-to-moment interaction. This may include how *specific relational messages* are being expressed by the client in the video clip. Optimal responses will include attempts to not only allude to the interpersonal tension, but make some attempt to repair that interpersonal issue.
- 4 Participant recognizes the other's interpersonal issue, and may discuss this further in more general terms (or discuss some secondary element of the other's issue or the relationship).
- 3 There may be more casual recognition of the interpersonal situation, but the response does not draw for further discussion of the issue or the relationship.
- 2 Participant addresses an issue to discuss that it is tangentially related to the interpersonal issue presented, but directs the discussion away from the present relationship situation.
- 1 Participant reacts to the interpersonal tension in a way that is nonproductive or in a way that likely exacerbates the rupture (e.g., responding negatively to a hostile client; responding to a controlling client with counter-control). Low scores also may be given when the participant fails to respond to the interpersonal issue involved in a way that indicates that the participant is avoiding the interpersonal issue or the relationship altogether.

Anexo V

Facilitative Interpersonal Skills – Self Report



Questionário FIS – Self Report

INSTRUÇÕES. Deu respostas verbais a breves situações retiradas de sessões reais, apresentadas em vídeo. Pensando em como foi para si realizar essa tarefa, indique por favor quanto cada afirmação reflecte a sua experiência com as respostas que deu aos clientes nestes vídeos. Assinale o número que melhor se adequa à sua resposta, de 1 (concordo muito) a 5 (discordo muito).

Não esqueça de preencher o seu nome e o número do vídeo. Obrigado!

Nome_

Nas minhas respostas aos vídeos eu acredito que...	Concordo Muito					Discordo Muito				
1. as minhas respostas foram empáticas e inferiram as emoções e experiências do cliente.....	1	2	3	4	5					
2. proporcionaram acolhimento e compreensão.....	1	2	3	4	5					
3. sugeriram ao cliente novas formas de pensar os seus problemas.....	1	2	3	4	5					
4. falei de forma a tornar claro ao cliente que eu queria colaborar com ele.....	1	2	3	4	5					
5. transmiti esperança, mesmo quando o cliente foi negativo ou hostil comigo.....	1	2	3	4	5					
6. senti que podia abordar o conflito interpessoal mencionado pelo cliente, e que consegui fazê-lo sem intensificar o conflito.....	1	2	3	4	5					
7. senti-me confortável quando falei e consegui usar um tom conversacional.....	1	2	3	4	5					
8. fui capaz de falar de transmitindo envolvimento emocional.....	1	2	3	4	5					

Anexo VI

Coeficiente de Correlação Intraclasse (CCI) (Output do SPSS)

	Correlação intraclasse ^b	Intervalo de Confiança 95%		Teste F com Valor True0			Sig
		Limite inferior	Limite superior	Valor	df1	df2	
Medidas únicas	,819 ^a	,795	,841	14,544	543	1086	<,001
Medidas médias	,931 ^c	,921	,941	14,544	543	1086	<,001

Modelo de efeitos mistos bidirecional em que os efeitos das pessoas são aleatórios e os das medidas são fixos.

- O estimador é o mesmo, esteja o efeito de interação presente ou não.
- Os coeficientes de correlação intraclasse tipo A que usam uma definição de concordância absoluta.
- Essa estimativa é calculada considerando que o efeito de interação esteja ausente, porque ele não pode ser estimado de outra forma.

Anexo VII

Teste de Normalidade dos Resíduos – FIS (Output do SPSS)

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	gl	Sig.	Estatística	gl	Sig.
Resíduo para M1	,147	17	,200*	,908	17	,093
Resíduo para M2	,135	17	,200*	,926	17	,186
Resíduo para M3	,150	17	,200*	,908	17	,094
Resíduo para M1.11	,166	17	,200*	,918	17	,134
Resíduo para M2.11	,133	17	,200*	,968	17	,787
Resíduo para M3.11	,194	17	,087	,903	17	,076
Resíduo para M1.12	,245	17	,008	,707	17	<,001
Resíduo para M2.12	,258	17	,004	,891	17	,049
Resíduo para M3.12	,209	17	,047	,836	17	,007
Resíduo para M1.13	,136	17	,200*	,975	17	,898
Resíduo para M2.13	,247	17	,007	,832	17	,006
Resíduo para M3.13	,215	17	,036	,911	17	,104
Resíduo para M1.14	,121	17	,200*	,961	17	,647
Resíduo para M2.14	,242	17	,009	,917	17	,132
Resíduo para M3.14	,126	17	,200*	,957	17	,578
Resíduo para M1.15	,251	17	,005	,680	17	<,001
Resíduo para M2.15	,166	17	,200*	,935	17	,261
Resíduo para M3.15	,179	17	,149	,903	17	,077
Resíduo para M1.16	,160	17	,200*	,895	17	,057
Resíduo para M2.16	,184	17	,128	,883	17	,035
Resíduo para M3.16	,125	17	,200*	,932	17	,235
Resíduo para M1.17	,187	17	,116	,925	17	,179
Resíduo para M2.17	,190	17	,106	,967	17	,759
Resíduo para M3.17	,256	17	,004	,767	17	<,001
Resíduo para M1.18	,123	17	,200*	,938	17	,300
Resíduo para M2.18	,107	17	,200*	,949	17	,444
Resíduo para M3.18	,280	17	<,001	,801	17	,002

*. Este é um limite inferior da significância verdadeira.

a. Correlação de Significância de Lilliefors

Anexo VIII

Teste de Normalidade dos Resíduos – FIS-SR (Output do SPSS)

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	gl	Sig.	Estatística	gl	Sig.
Resíduo para Auto.M1	,124	17	,200 [*]	,957	17	,578
Resíduo para Auto.M2	,137	17	,200 [*]	,983	17	,982
Resíduo para Auto.M3	,182	17	,138	,927	17	,195
Resíduo para Auto.M1.I1	,214	17	,037	,920	17	,148
Resíduo para Auto.M2.I1	,227	17	,020	,939	17	,304
Resíduo para Auto.M3.I1	,245	17	,008	,785	17	,001
Resíduo para Auto.M1.I2	,101	17	,200 [*]	,974	17	,882
Resíduo para Auto.M2.I2	,131	17	,200 [*]	,954	17	,525
Resíduo para Auto.M3.I2	,164	17	,200 [*]	,938	17	,293
Resíduo para Auto.M1.I3	,100	17	,200 [*]	,977	17	,925
Resíduo para Auto.M2.I3	,116	17	,200 [*]	,938	17	,300
Resíduo para Auto.M3.I3	,151	17	,200 [*]	,965	17	,726
Resíduo para Auto.M1.I4	,162	17	,200 [*]	,963	17	,690
Resíduo para Auto.M2.I4	,095	17	,200 [*]	,978	17	,941
Resíduo para Auto.M3.I4	,233	17	,015	,846	17	,009
Resíduo para Auto.M1.I5	,189	17	,109	,954	17	,530
Resíduo para Auto.M2.I5	,147	17	,200 [*]	,931	17	,225
Resíduo para Auto.M3.I5	,267	17	,002	,863	17	,017
Resíduo para Auto.M1.I6	,111	17	,200 [*]	,985	17	,988
Resíduo para Auto.M2.I6	,175	17	,176	,919	17	,145
Resíduo para Auto.M3.I6	,209	17	,048	,889	17	,045
Resíduo para Auto.M1.I7	,124	17	,200 [*]	,958	17	,587
Resíduo para Auto.M2.I7	,170	17	,200 [*]	,976	17	,911
Resíduo para Auto.M3.I7	,220	17	,028	,846	17	,009
Resíduo para Auto.M1.I8	,154	17	,200 [*]	,937	17	,288
Resíduo para Auto.M2.I8	,121	17	,200 [*]	,954	17	,514
Resíduo para Auto.M3.I8	,216	17	,034	,893	17	,052

*. Este é um limite inferior da significância verdadeira.

a. Correlação de Significância de Lilliefors

Anexo IX

ANOVA Medidas Repetidas – FIS Global (Outputs do SPSS)

Teste de esfericidade de Mauchly^a

Medida: MEASURE_1

Efeito dentre-sujeitos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Momento	,317	17,222	2	<,001	,594	,615	,500

Testa a hipótese nula para a qual a matriz de covariâncias de erro das variáveis transformadas ortonormalizadas é proporcional em relação a uma matriz identidade.

a. Design: Intercepto

Design Dentre-Sujeitos: Momento

b. Pode ser usado para ajustar os graus de liberdade dos testes de significância dentro da média. Os testes corrigidos são exibidos na tabela Testes de efeitos dentre-sujeitos.

Testes de efeitos dentre-sujeitos

Medida: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	F	Sig.	Eta parcial quadrado
Momento	Esfericidade considerada	4,462	2	2,231	72,733	<,001	,820
	Greenhouse-Geisser	4,462	1,189	3,755	72,733	<,001	,820
	Huynh-Feldt	4,462	1,229	3,631	72,733	<,001	,820
	Limite inferior	4,462	1,000	4,462	72,733	<,001	,820
Erro(Momento)	Esfericidade considerada	,982	32	,031			
	Greenhouse-Geisser	,982	19,016	,052			
	Huynh-Feldt	,982	19,666	,050			
	Limite inferior	,982	16,000	,061			

Comparações por Método Pairwise

Medida: MEASURE_1

(I) Momento	(J) Momento	Diferença média (I-J)	Estatística do teste Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	-,282 [*]	,041	<,001	-,391	-,173
	3	-,719 [*]	,081	<,001	-,935	-,503
2	1	,282 [*]	,041	<,001	,173	,391
	3	-,437 [*]	,051	<,001	-,574	-,300
3	1	,719 [*]	,081	<,001	,503	,935
	2	,437 [*]	,051	<,001	,300	,574

Baseado em médias marginais estimadas

*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: Bonferroni.

Anexo X

Anova Medidas Repetidas – FIS – Itens 1 e 4 (Outputs do SPSS)

X.1 — ANOVA Medidas Repetidas – FIS – Item 1 (i1)

Teste de esfericidade de Mauchly^a

Medida: MEASURE_1

Efeito dentre-sujeitos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Momento	,515	9,958	2	,007	,673	,713	,500

Testa a hipótese nula para a qual a matriz de covariâncias de erro das variáveis transformadas ortonormalizadas é proporcional em relação a uma matriz identidade.

a. Design: Intercepto

Design Dentre-Sujeitos: Momento

b. Pode ser usado para ajustar os graus de liberdade dos testes de significância dentro da média. Os testes corrigidos são exibidos na tabela Testes de efeitos dentre-sujeitos.

Testes de efeitos dentre-sujeitos

Medida: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	F	Sig.	Eta parcial quadrado
Momento	Esfericidade considerada	1,904	2	,952	37,243	<,001	,699
	Greenhouse-Geisser	1,904	1,347	1,414	37,243	<,001	,699
	Huynh-Feldt	1,904	1,426	1,335	37,243	<,001	,699
	Limite inferior	1,904	1,000	1,904	37,243	<,001	,699
Erro(Momento)	Esfericidade considerada	,818	32	,026			
	Greenhouse-Geisser	,818	21,547	,038			
	Huynh-Feldt	,818	22,814	,036			
	Limite inferior	,818	16,000	,051			

Comparações por Método Pairwise

Medida: MEASURE_1

(I) Momento	(J) Momento	Diferença média (I-J)	Estatística do teste Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	-,209 [*]	,030	<,001	-,289	-,128
	3	-,472 [*]	,063	<,001	-,642	-,303
2	1	,209 [*]	,030	<,001	,128	,289
	3	-,264 [*]	,064	,002	-,434	-,093
3	1	,472 [*]	,063	<,001	,303	,642
	2	,264 [*]	,064	,002	,093	,434

Baseado em médias marginais estimadas

*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: Bonferroni.

X.2 — ANOVA Medidas Repetidas – FIS – Item 4 (i4)

Teste de esfericidade de Mauchly^a

Medida: MEASURE_1

Efeito dentre-sujeitos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Momento	,732	4,677	2	,096	,789	,860	,500

Testa a hipótese nula para a qual a matriz de covariâncias de erro das variáveis transformadas ortonormalizadas é proporcional em relação a uma matriz identidade.

a. Design: Intercepto

Design Dentre-Sujeitos: Momento

b. Pode ser usado para ajustar os graus de liberdade dos testes de significância dentro da média. Os testes corrigidos são exibidos na tabela Testes de efeitos dentre-sujeitos.

Testes de efeitos dentre-sujeitos

Medida: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	F	Sig.	Eta parcial quadrado
Momento	Esfericidade considerada	2,327	2	1,164	41,017	<,001	,719
	Greenhouse-Geisser	2,327	1,577	1,475	41,017	<,001	,719
	Huynh-Feldt	2,327	1,721	1,353	41,017	<,001	,719
	Limite inferior	2,327	1,000	2,327	41,017	<,001	,719
Erro(Momento)	Esfericidade considerada	,908	32	,028			
	Greenhouse-Geisser	,908	25,240	,036			
	Huynh-Feldt	,908	27,531	,033			
	Limite inferior	,908	16,000	,057			

Comparações por Método Pairwise

Medida: MEASURE_1

(I) Momento	(J) Momento	Diferença média (I-J)	Estatística do teste Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	-,359 [*]	,048	<,001	-,488	-,230
	3	-,509 [*]	,071	<,001	-,699	-,319
2	1	,359 [*]	,048	<,001	,230	,488
	3	-,150 [*]	,051	,029	-,287	-,014
3	1	,509 [*]	,071	<,001	,319	,699
	2	,150 [*]	,051	,029	,014	,287

Baseado em médias marginais estimadas

*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: Bonferroni.

Anexo XI

Friedman – FIS – Itens 2, 3 e 5 a 8 (Outputs do SPSS)

XI.1 — Friedman – FIS – Item 2 (i2)

Estadísticas de teste^a

N	17
Qui-quadrado	30,406
df	2
Significância Sig.	<,001

a. Teste Friedman

XI.2 — Friedman – FIS – Item 3 (i3)

Estadísticas de teste^a

N	17
Qui-quadrado	34,000
df	2
Significância Sig.	<,001

a. Teste Friedman

XI.3 — Friedman – FIS – Item 5 (i5)

Estadísticas de teste^a

N	17
Qui-quadrado	27,882
df	2
Significância Sig.	<,001

a. Teste Friedman

XI.4 — Friedman – FIS – Item 6 (i6)

Estadísticas de teste^a

N	17
Qui-quadrado	27,672
df	2
Significância Sig.	<,001

a. Teste Friedman

XI.5 — Friedman – FIS – Item 7 (i7)

Estadísticas de teste^a

N	17
Qui-quadrado	27,882
df	2
Significância Sig.	<,001

a. Teste Friedman

XI.6 — Friedman – FIS – Item 8 (i8)

Estatísticas de teste^a

N	17
Qui-quadrado	29,059
df	2
Significância Sig.	<,001

a. Teste Friedman

Anexo XII

Wilcoxon – FIS – Itens 2, 3 e 5 a 8 (Outputs do SPSS)

XII.1 — Wilcoxon – FIS – Item 2 (i2)

Estatísticas de teste^a

	M2.I2 - M1.I2	M3.I2 - M1.I2	M3.I2 - M2.I2
Z	-3,519 ^b	-3,622 ^b	-2,983 ^b
Significância Sig. (2 extremidades)	<,001	<,001	,003

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XII.2 — Wilcoxon – FIS – Item 3 (i3)

Estatísticas de teste^a

	M2.I3 - M1.I3	M3.I3 - M1.I3	M3.I3 - M2.I3
Z	-3,622 ^b	-3,621 ^b	-3,622 ^b
Significância Sig. (2 extremidades)	<,001	<,001	<,001

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XII.3 — Wilcoxon – FIS – Item 5 (i5)

Estatísticas de teste^a

	M2.I5 - M1.I5	M3.I5 - M1.I5	M3.I5 - M2.I5
Z	-2,984 ^b	-3,622 ^b	-3,622 ^b
Significância Sig. (2 extremidades)	,003	<,001	<,001

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XII.4 — Wilcoxon – FIS – Item 6 (i6)

Estatísticas de teste^a

	M2.I6 - M1.I6	M3.I6 - M1.I6	M3.I6 - M2.I6
Z	-2,911 ^b	-3,516 ^b	-3,622 ^b
Significância Sig. (2 extremidades)	,004	<,001	<,001

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XII.5 — Wilcoxon – FIS – Item 7 (i7)

Estadísticas de teste^a

	M2.17 - M1.17	M3.17 - M1.17	M3.17 - M2.17
Z	-2,866 ^b	-3,621 ^b	-3,623 ^b
Significância Sig. (2 extremidades)	,004	<,001	<,001

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XII.6 — Wilcoxon – FIS – Item 8 (i8)

Estadísticas de teste^a

	M2.18 - M1.18	M3.18 - M1.18	M3.18 - M2.18
Z	-3,243 ^b	-3,621 ^b	-3,622 ^b
Significância Sig. (2 extremidades)	,001	<,001	<,001

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

Anexo XIII

ANOVA Medidas Repetidas – FIS-SR Global (Outputs do SPSS)

Teste de esfericidade de Mauchly^a

Medida: MEASURE_1

Efeito dentre-sujeitos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Momento	,813	3,108	2	,211	,842	,931	,500

Testa a hipótese nula para a qual a matriz de covariâncias de erro das variáveis transformadas ortonormalizadas é proporcional em relação a uma matriz identidade.

a. Design: Intercepto

Design Dentre-Sujeitos: Momento

b. Pode ser usado para ajustar os graus de liberdade dos testes de significância dentro da média. Os testes corrigidos são exibidos na tabela Testes de efeitos dentre-sujeitos.

Testes de efeitos dentre-sujeitos

Medida: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	F	Sig.	Eta parcial quadrado
	Greenhouse-Geisser	2,188	1,685	1,299	4,841	,020	,232
	Huynh-Feldt	2,188	1,861	1,176	4,841	,017	,232
	Limite inferior	2,188	1,000	2,188	4,841	,043	,232
Erro(Momento)	Esfericidade considerada	7,232	32	,226			
	Greenhouse-Geisser	7,232	26,956	,268			
	Huynh-Feldt	7,232	29,776	,243			
	Limite inferior	7,232	16,000	,452			

Comparações por Método Pairwise

Medida: MEASURE_1

(I) Momento	(J) Momento	Diferença média (I-J)	Estatística do teste Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	-,410	,175	,098	-,879	,058
	3	-,464*	,123	,005	-,794	-,134
2	1	,410	,175	,098	-,058	,879
	3	-,054	,184	1,000	-,545	,438
3	1	,464*	,123	,005	,134	,794
	2	,054	,184	1,000	-,438	,545

Baseado em médias marginais estimadas

*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: Bonferroni.

Anexo XIV

ANOVA Medidas Repetidas – FIS-SR – Itens 2, 3 e 8 (Outputs do SPSS)

XIV.1 — ANOVA Medidas Repetidas – FIS-SR – Item 2 (i2)

Teste de esfericidade de Mauchly^a

Medida: MEASURE_1

Efeito dentre-sujeitos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Momento	,952	,731	2	,694	,955	1,000	,500

Testa a hipótese nula para a qual a matriz de covariâncias de erro das variáveis transformadas ortonormalizadas é proporcional em relação a uma matriz identidade.

a. Design: Intercepto

Design Dentre-Sujeitos: Momento

b. Pode ser usado para ajustar os graus de liberdade dos testes de significância dentro da média. Os testes corrigidos são exibidos na tabela Testes de efeitos dentre-sujeitos.

Testes de efeitos dentre-sujeitos

Medida: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	F	Sig.	Eta parcial quadrado
Momento	Esfericidade considerada	1,208	2	,604	2,799	,076	,149
	Greenhouse-Geisser	1,208	1,909	,633	2,799	,079	,149
	Huynh-Feldt	1,208	2,000	,604	2,799	,076	,149
	Limite inferior	1,208	1,000	1,208	2,799	,114	,149
Erro(Momento)	Esfericidade considerada	6,907	32	,216			
	Greenhouse-Geisser	6,907	30,547	,226			
	Huynh-Feldt	6,907	32,000	,216			
	Limite inferior	6,907	16,000	,432			

Comparações por Método Pairwise

Medida: MEASURE_1

(I) Momento	(J) Momento	Diferença média (I-J)	Estatística do teste Padrão	Sig. ^a	95% Intervalo de Confiança para Diferença ^a	
					Limite inferior	Limite superior
1	2	-,261	,165	,398	-,703	,180
	3	-,366	,141	,059	-,744	,012
2	1	,261	,165	,398	-,180	,703
	3	-,105	,170	1,000	-,560	,351
3	1	,366	,141	,059	-,012	,744
	2	,105	,170	1,000	-,351	,560

Baseado em médias marginais estimadas

a. Ajustamento para diversas comparações: Bonferroni.

XIV.2 — ANOVA Medidas Repetidas – FIS-SR – Item 3 (i3)

Teste de esfericidade de Mauchly^a

Medida: MEASURE_1

Efeito dentre-sujeitos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Momento	,869	2,099	2	,350	,884	,986	,500

Testa a hipótese nula para a qual a matriz de covariâncias de erro das variáveis transformadas ortonormalizadas é proporcional em relação a uma matriz identidade.

a. Design: Intercepto

Design Dentre-Sujeitos: Momento

b. Pode ser usado para ajustar os graus de liberdade dos testes de significância dentro da média. Os testes corrigidos são exibidos na tabela Testes de efeitos dentre-sujeitos.

Testes de efeitos dentre-sujeitos

Medida: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	F	Sig.	Eta parcial quadrado
Momento	Esfericidade considerada	2,023	2	1,012	4,433	,020	,217
	Greenhouse-Geisser	2,023	1,769	1,144	4,433	,025	,217
	Huynh-Feldt	2,023	1,973	1,026	4,433	,020	,217
	Limite inferior	2,023	1,000	2,023	4,433	,051	,217
Erro(Momento)	Esfericidade considerada	7,302	32	,228			
	Greenhouse-Geisser	7,302	28,303	,258			
	Huynh-Feldt	7,302	31,562	,231			
	Limite inferior	7,302	16,000	,456			

Comparações por Método Pairwise

Medida: MEASURE_1

(I) Momento	(J) Momento	Diferença média (I-J)	Estatística do teste Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	-,438	,184	,090	-,929	,054
	3	-,405*	,132	,022	-,759	-,051
2	1	,438	,184	,090	-,054	,929
	3	,033	,171	1,000	-,424	,490
3	1	,405*	,132	,022	,051	,759
	2	-,033	,171	1,000	-,490	,424

Baseado em médias marginais estimadas

*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: Bonferroni.

XIV.3 — ANOVA Medidas Repetidas – FIS-SR – Item 8 (i8)

Teste de esfericidade de Mauchly^a

Medida: MEASURE_1

Efeito dentre-sujeitos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Momento	,719	4,952	2	,084	,781	,850	,500

Testa a hipótese nula para a qual a matriz de covariâncias de erro das variáveis transformadas ortonormalizadas é proporcional em relação a uma matriz identidade.

a. Design: Intercepto

Design Dentre-Sujeitos: Momento

b. Pode ser usado para ajustar os graus de liberdade dos testes de significância dentro da média. Os testes corrigidos são exibidos na tabela Testes de efeitos dentre-sujeitos.

Testes de efeitos dentre-sujeitos

Medida: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	F	Sig.	Eta parcial quadrado
Momento	Esfericidade considerada	2,321	2	1,161	4,214	,024	,208
	Greenhouse-Geisser	2,321	1,561	1,487	4,214	,035	,208
	Huynh-Feldt	2,321	1,699	1,366	4,214	,031	,208
	Limite inferior	2,321	1,000	2,321	4,214	,057	,208
Erro(Momento)	Esfericidade considerada	8,813	32	,275			
	Greenhouse-Geisser	8,813	24,977	,353			
	Huynh-Feldt	8,813	27,191	,324			
	Limite inferior	8,813	16,000	,551			

Comparações por Método Pairwise

Medida: MEASURE_1

(I) Momento	(J) Momento	Diferença média (I-J)	Estatística do teste Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	-,222	,167	,607	-,669	,225
	3	-,521 [*]	,142	,006	-,901	-,140
2	1	,222	,167	,607	-,225	,669
	3	-,298	,221	,589	-,890	,293
3	1	,521 [*]	,142	,006	,140	,901
	2	,298	,221	,589	-,293	,890

Baseado em médias marginais estimadas

*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: Bonferroni.

Anexo XV

Friedman – FIS-SR – Itens 1, 4, 5, 6, e 7 (Outputs do SPSS)

XV.1 — Friedman – FIS-SR – Item 1 (i1)

Estatísticas de teste^a

N	17
Qui-quadrado	9,406
df	2
Significância Sig.	,009

a. Teste Friedman

XV.2 — Friedman – FIS-SR – Item 4 (i4)

Estatísticas de teste^a

N	17
Qui-quadrado	8,885
df	2
Significância Sig.	,012

a. Teste Friedman

XV.3 — Friedman – FIS-SR – Item 5 (i5)

Estatísticas de teste^a

N	17
Qui-quadrado	8,328
df	2
Significância Sig.	,016

a. Teste Friedman

XV.4 — Friedman – FIS-SR – Item 6 (i6)

Estatísticas de teste^a

N	17
Qui-quadrado	7,394
df	2
Significância Sig.	,025

a. Teste Friedman

XV.5 — Friedman – FIS-SR – Item 7 (i7)

Estatísticas de teste^a

N	17
Qui-quadrado	12,317
df	2
Significância Sig.	,002

a. Teste Friedman

Anexo XVI

Wilcoxon – FIS-SR – Itens 1, 4, 5, 6, e 7 (Outputs do SPSS)

XVI.1 — Wilcoxon – FIS-SR – Item 1 (i1)

Estadísticas de teste^a

	Auto.M2.I1 - Auto.M1.I1	Auto.M3.I1 - Auto.M1.I1	Auto.M3.I1 - Auto.M2.I1
Z	-1,579 ^b	-2,612 ^b	-1,606 ^b
Significância Sig. (2 extremidades)	,114	,009	,108

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XVI.2 — Wilcoxon – FIS-SR – Item 4 (i4)

Estadísticas de teste^a

	Auto.M2.I4 - Auto.M1.I4	Auto.M3.I4 - Auto.M1.I4	Auto.M3.I4 - Auto.M2.I4
Z	-2,123 ^b	-2,449 ^b	-,280 ^c
Significância Sig. (2 extremidades)	,034	,014	,779

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

c. Com base em postos positivos.

XVI.3 — Wilcoxon – FIS-SR – Item 5 (i5)

Estadísticas de teste^a

	Auto.M2.I5 - Auto.M1.I5	Auto.M3.I5 - Auto.M1.I5	Auto.M3.I5 - Auto.M2.I5
Z	-2,324 ^b	-2,681 ^b	-,026 ^c
Significância Sig. (2 extremidades)	,020	,007	,979

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

c. Com base em postos positivos.

XVI.4 — Wilcoxon – FIS-SR – Item 6 (i6)

Estatísticas de teste^a

	Auto.M2.I6 - Auto.M1.I6	Auto.M3.I6 - Auto.M1.I6	Auto.M3.I6 - Auto.M2.I6
Z	-2,440 ^b	-2,465 ^b	-,114 ^c
Significância Sig. (2 extremidades)	,015	,014	,909

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

c. Com base em postos positivos.

XVI.5 — Wilcoxon – FIS-SR – Item 7 (i7)**Estatísticas de teste^a**

	Auto.M2.I7 - Auto.M1.I7	Auto.M3.I7 - Auto.M1.I7	Auto.M3.I7 - Auto.M2.I7
Z	-2,613 ^b	-1,295 ^b	-1,427 ^c
Significância Sig. (2 extremidades)	,009	,195	,154

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

c. Com base em postos positivos.

Anexo XVII

ANOVA Medidas Repetidas – 2 Fatores – FIS e FIS-SR Globais (Outputs do SPSS)

Teste de esfericidade de Mauchly^a

Medida: MEASURE_1

Efeito dentre-sujeitos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Momento	,981	,291	2	,864	,981	1,000	,500
TipodeAvaliação	1,000	,000	0	.	1,000	1,000	1,000
Momento * TipodeAvaliação	,733	4,650	2	,098	,790	,861	,500

Testa a hipótese nula para a qual a matriz de covariâncias de erro das variáveis transformadas ortonormalizadas é proporcional em relação a uma matriz identidade.

a. Design: Intercepto

Design Dentre-Sujeitos: Momento + TipodeAvaliação + Momento * TipodeAvaliação

b. Pode ser usado para ajustar os graus de liberdade dos testes de significância dentro da média. Os testes corrigidos são exibidos na tabela Testes de efeitos dentre-sujeitos.

Testes de efeitos dentre-sujeitos

Medida: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	F	Sig.	Eta parcial quadrado
Momento	Esfericidade considerada	6,004	2	3,002	19,594	<,001	,550
	Greenhouse-Geisser	6,004	1,962	3,060	19,594	<,001	,550
	Huynh-Feldt	6,004	2,000	3,002	19,594	<,001	,550
	Limite inferior	6,004	1,000	6,004	19,594	<,001	,550
Erro(Momento)	Esfericidade considerada	4,903	32	,153			
	Greenhouse-Geisser	4,903	31,396	,156			
	Huynh-Feldt	4,903	32,000	,153			
	Limite inferior	4,903	16,000	,306			
TipodeAvaliação	Esfericidade considerada	2,135	1	2,135	11,434	,004	,417
	Greenhouse-Geisser	2,135	1,000	2,135	11,434	,004	,417
	Huynh-Feldt	2,135	1,000	2,135	11,434	,004	,417
	Limite inferior	2,135	1,000	2,135	11,434	,004	,417
Erro(TipodeAvaliação)	Esfericidade considerada	2,987	16	,187			
	Greenhouse-Geisser	2,987	16,000	,187			
	Huynh-Feldt	2,987	16,000	,187			
	Limite inferior	2,987	16,000	,187			
Momento * TipodeAvaliação	Esfericidade considerada	,647	2	,323	3,126	,058	,163
	Greenhouse-Geisser	,647	1,579	,410	3,126	,072	,163
	Huynh-Feldt	,647	1,723	,375	3,126	,066	,163
	Limite inferior	,647	1,000	,647	3,126	,096	,163
Erro (Momento*TipodeAvaliação)	Esfericidade considerada	3,311	32	,103			
	Greenhouse-Geisser	3,311	25,266	,131			
	Huynh-Feldt	3,311	27,565	,120			
	Limite inferior	3,311	16,000	,207			

Comparações por Método Pairwise

Medida: MEASURE_1

(I) TipodeAvaliação	(J) TipodeAvaliação	Diferença média (I-J)	Estatística do teste Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	-,289 [*]	,086	,004	-,471	-,108
2	1	,289 [*]	,086	,004	,108	,471

Baseado em médias marginais estimadas

*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: Bonferroni.

Anexo XVIII

Wilcoxon – Comparação FIS e FIS-SR – Itens 1 a 8 (Outputs do SPSS)

XVIII.1 — Wilcoxon – Comparação FIS e FIS-SR – Item 1 (i1)

Estadísticas de teste^a

	Auto.M1.I1 - M1.I1	Auto.M2.I1 - M2.I1	Auto.M3.I1 - M3.I1
Z	-,829 ^b	-,994 ^b	-2,107 ^b
Significância Sig. (2 extremidades)	,407	,320	,035

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XVIII.2 — Wilcoxon – Comparação FIS e FIS-SR – Item 2 (i2)

Estadísticas de teste^a

	Auto.M1.I2 - M1.I2	Auto.M2.I2 - M2.I2	Auto.M3.I2 - M3.I2
Z	-1,009 ^b	-,923 ^b	-,403 ^b
Significância Sig. (2 extremidades)	,313	,356	,687

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XVIII.3 — Wilcoxon – Comparação FIS e FIS-SR – Item 3 (i3)

Estadísticas de teste^a

	Auto.M1.I3 - M1.I3	Auto.M2.I3 - M2.I3	Auto.M3.I3 - M3.I3
Z	-1,254 ^b	-,852 ^b	-1,231 ^c
Significância Sig. (2 extremidades)	,210	,394	,218

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

c. Com base em postos positivos.

XVIII.4 — Wilcoxon – Comparação FIS e FIS-SR – Item 4 (i4)

Estatísticas de teste^a

	Auto.M1.I4 - M1.I4	Auto.M2.I4 - M2.I4	Auto.M3.I4 - M3.I4
Z	-3,197 ^b	-2,663 ^b	-2,249 ^b
Significância Sig. (2 extremidades)	,001	,008	,025

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XVIII.5 — Wilcoxon – Comparação FIS e FIS-SR – Item 5 (i5)**Estatísticas de teste^a**

	Auto.M1.I5 - M1.I5	Auto.M2.I5 - M2.I5	Auto.M3.I5 - M3.I5
Z	-2,581 ^b	-2,722 ^b	-,362 ^b
Significância Sig. (2 extremidades)	,010	,006	,717

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XVIII.6 — Wilcoxon – Comparação FIS e FIS-SR – Item 6 (i6)**Estatísticas de teste^a**

	Auto.M1.I6 - M1.I6	Auto.M2.I6 - M2.I6	Auto.M3.I6 - M3.I6
Z	-2,367 ^b	-2,557 ^b	-2,069 ^b
Significância Sig. (2 extremidades)	,018	,011	,039

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

XVIII.7 — Wilcoxon – Comparação FIS e FIS-SR – Item 7 (i7)

Estatísticas de teste^a

	Auto.M1.17 - M1.17	Auto.M2.17 - M2.17	Auto.M3.17 - M3.17
Z	-2,770 ^b	-2,582 ^b	-1,516 ^c
Significância Sig. (2 extremidades)	,006	,010	,129

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

c. Com base em postos positivos.

XVIII.8 — Wilcoxon – Comparação FIS e FIS-SR – Item 8 (i8)**Estatísticas de teste^a**

	Auto.M1.18 - M1.18	Auto.M2.18 - M2.18	Auto.M3.18 - M3.18
Z	-3,528 ^b	-2,320 ^b	-,853 ^b
Significância Sig. (2 extremidades)	<,001	,020	,394

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.