

Artigos de pesquisa

Educação baseada em telessimulação: validação de um programa instrucional para o desenvolvimento da competência em diálogo socrático em terapia cognitivo-comportamental

Telesimulation-based education: validation of an instructional program for developing socratic questioning skills in cognitive behavioral therapy

Roberta Gonçalves Joaquim^{1*} , Janaína Bianca Barletta² , Fernando Alves de Lima Franco¹ 

¹Faculdade de Ciências da Saúde Albert Einstein (FICSAE), Programa de Pós-graduação de Ensino em Saúde, São Paulo, SP, Brasil

²Associação de Ensino e Supervisão Baseados em Evidências (AESBE), São Paulo, SP, Brasil

COMO CITAR: JOAQUIM, R. G.; BARLETTA, J. B.; FRANCO, F. A. L. Educação baseada em telessimulação: validação de um programa instrucional para o desenvolvimento da competência em diálogo socrático em terapia cognitivo-comportamental. *Revista Ibero-Americana de Estudos em Educação*, Araraquara, v. 21, e19212, 2026. e-ISSN: 1982-5587. DOI: <https://doi.org/10.21723/riaee.v21i00.1921201>

Resumo

A telessimulação integra simulação clínica e educação mediada por tecnologia, ampliando o acesso ao treino supervisionado e ao ensino por competências, sobretudo quando se requer prática deliberada e *feedback* estruturado. Quando direcionamos essa abordagem à formação de psicólogos, especialmente em Terapia Cognitivo-Comportamental (TCC), na qual o diálogo socrático (DS) depende de micro-habilidades e devolutivas qualificadas, a telessimulação parece promissora. Este estudo metodológico teve como objetivo elaborar e validar o conteúdo de um programa instrucional de telessimulação para desenvolver DS em terapeutas em formação continuada. O programa incluiu guia de sistematização, avaliação de conhecimento, cenários síncronos com paciente padronizado e *debriefing* estruturado, sendo avaliado por juízes em duas rodadas quanto à clareza, pertinência e aplicabilidade. Após revisões, os itens reavaliados alcançaram concordância plena (RVC/CVR=1,0). As limitações incluem evidências restritas à validade de conteúdo, sendo assim recomenda-se estudo piloto e pesquisas sobre efetividade, transferência para a prática e implementação em diferentes contextos.

Palavras-chave: educação em saúde; educação continuada; treinamento por simulação; terapia cognitivo-comportamental.

Abstract

Telesimulation integrates clinical simulation and technology-mediated education, expanding access to supervised training and skill-based teaching, especially when deliberate practice and structured feedback are required. When we direct this approach to the training of psychologists, especially in Cognitive Behavioral Therapy (CBT), in which Socratic questioning (SQ) depends on micro-skills and qualified feedback, telesimulation seems promising. This methodological study aimed to develop and validate the content of a telesimulation instructional program to develop SQ in therapists in continuing education. The program included a systematization guide, knowledge assessment, synchronous scenarios with a standardized patient, and structured debriefing, which was evaluated by judges in two rounds for clarity, relevance, and applicability. After revisions, the re-evaluated items achieved full agreement (RVC/CVR=1.0). Limitations include evidence restricted to content validity, so a pilot study and research on effectiveness, transfer to practice, and implementation in different contexts is recommended.

Keywords: health education; continuing education; simulation training; cognitive behavioral therapy.

INTRODUÇÃO

O cenário contemporâneo, marcado pelas transformações do contexto pós-pandêmico, evidenciou a distribuição geográfica não apenas como um desafio logístico, mas como um

*Autor correspondente:

robertagjoaquim@gmail.com

Submetido: Abril 16, 2024

Revisado: Janeiro 26, 2026

Aprovado: Janeiro 26, 2026

Fonte de financiamento: nada a declarar.

Conflitos de interesse: Não há conflitos de interesse.

Aprovação do comitê de ética: Esse estudo foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos da Faculdade Israelita de Ciências da Saúde Albert Einstein (CEP-HIAE), conforme o Parecer consubstanciado nº 5.934.557 e todos os participantes assinaram o termo de consentimento livre e esclarecido (TCLE).

Disponibilidade de dados: Os dados estarão disponíveis mediante pedido aos autores.

Trabalho realizado na Faculdade de Ciências da Saúde Albert Einstein (FICSAE), São Paulo, SP, Brasil.



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/>), que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

fator social que impacta a equidade no acesso à educação continuada. No âmbito da educação em saúde, a distância dos grandes centros pode tornar-se barreira ao desenvolvimento e à manutenção de competências complexas, exigindo respostas inovadoras para ampliar a capilaridade do ensino de excelência. Nesse contexto, as inovações digitais têm se consolidado como instrumentos para mitigar assimetrias e conectar profissionais dispersos territorialmente a oportunidades de treino estruturado (Hayden et al., 2018; Naik et al., 2020).

Entre essas estratégias, a telessimulação tem ganhado destaque por combinar princípios da simulação realística com recursos de telecomunicação, viabilizando experiências de aprendizagem remotas e/ou síncronas, com aprendizes e facilitadores em locais distintos (Thomas et al., 2021; Hayden et al., 2018). Ao reduzir barreiras geográficas e logísticas, essa modalidade amplia o acesso a práticas supervisionadas e favorece a continuidade do treinamento, especialmente quando se requerem atividades observáveis e passíveis de devolutiva (Thomas et al., 2021).

Do ponto de vista pedagógico, a telessimulação se apoia na teoria da aprendizagem experiencial de Kolb (1984), na qual o aprendizado se organiza em um ciclo que parte da experiência concreta, avança para a observação reflexiva e a conceituação abstrata, e se consolida na experimentação ativa. Nesse desenho, o *debriefing* ocupa papel central ao sustentar uma reflexão orientada sobre decisões, comportamentos e raciocínios mobilizados durante e após a experiência, fortalecendo a aprendizagem pela análise do desempenho e favorecendo a transferência para situações futuras (Watts et al., 2021). Para que esse potencial formativo se concretize, contudo, é necessário rigor instrucional: diretrizes internacionais, descritas pela *International Nursing Association for Clinical Simulation and Learning* (INACSL), recomendam planejamento intencional, com alinhamento entre objetivos, recursos, fidelidade do cenário e estratégias de facilitação, de modo a maximizar resultados educacionais (Watts et al., 2021).

No campo da Psicologia, essa discussão se insere em uma trajetória formativa que se inicia na graduação, mas se consolida de modo decisivo após o diploma, em cursos de especialização, programas de formação e supervisão clínica continuada. Estudos sobre a formação em Psicologia no Brasil apontam desafios na articulação entre teoria, prática e desenvolvimento de competências para o exercício profissional (Rudá; Coutinho; Almeida Filho, 2019). Em paralelo, a formação em Terapia Cognitivo-Comportamental (TCC) na graduação é descrita como heterogênea e, por vezes, insuficiente para sustentar o domínio aplicado de competências terapêuticas, o que contribui para a busca de complementação formativa em contextos de pós-graduação e educação continuada (Neufeld et al., 2018).

Esse movimento é especialmente relevante quando se considera que intervenções baseadas em evidências exigem mais do que conhecimento declarativo: requerem desempenho observável, *feedback* qualificado e oportunidades repetidas de prática deliberada. Revisões sobre treinamento de terapeutas indicam limitações de formatos exclusivamente expositivos e apontam melhores resultados quando o ensino integra treino ativo, *feedback* e suporte/consulta (Frank; Becker-Haimes; Kendall, 2020). Na formação em TCC, essa demanda se intensifica, pois competências centrais dependem de micro-habilidades clínicas refinadas e treináveis (Beck, 2021; Roth; Pilling, 2008). Nesse enquadre, a descoberta guiada organiza-se como modo de condução que apoia o cliente a examinar significados e alternativas a partir de suas próprias experiências, e o diálogo socrático (DS) constitui uma de suas expressões centrais, não como “roteiro de perguntas”, mas como postura técnica de curiosidade, colaboração, precisão e sensibilidade ao contexto (Kazantzis et al., 2018; Padesky, 2019; Overholser, 2010). Estudos com clínicos em formação apontam que a descoberta guiada figura entre as competências mais difíceis de dominar, com obstáculos relacionados a condução apressada, excesso de diretividade e dificuldade de equilibrar direção clínica e colaboração (Waltman et al., 2017). Nesse cenário, a telessimulação surge como alternativa promissora para organizar experiências de treino com maior padronização e observação direta; contudo, a efetividade educacional também depende de autenticidade e desenho cuidadoso do cenário, incluindo treinamento do paciente padronizado (SP) (Kühne et al., 2018).

Apesar desse conjunto de recomendações, ainda são escassas propostas aplicadas que sistematizem o uso da telessimulação, com SP e *debriefing* estruturado, para o ensino de competências terapêuticas específicas na formação em TCC, especialmente em contextos de pós-graduação e educação continuada. Assim, este estudo teve como objetivo elaborar e

validar, quanto ao conteúdo, um programa de treinamento baseado em telessimulação para promover o desenvolvimento e o aprimoramento da competência terapêutica de diálogo socrático em terapeutas em formação em TCC.

MÉTODOS

O estudo adota uma abordagem metodológica mista, combinando elementos quantitativos e qualitativos. Dividido em duas etapas distintas, a primeira focou na elaboração do programa de treinamento e do instrumento de avaliação do conhecimento, enquanto a segunda concentrou-se na validação de conteúdo por juízes especialistas. A análise visou avaliar a qualidade do conteúdo do treinamento e da avaliação e conhecimento, considerando sua abrangência, clareza, pertinência, objetividade e praticidade, conforme orientações metodológicas previamente estabelecidas (Alexandre; Coluci, 2011).

Considerando o objetivo educacional do treinamento e as orientações de boas práticas em simulação sugerido por Watts et al. (2021), adotou-se o *Debriefing for Meaningful Learning* (DML) como modelo de *debriefing*, dada a necessidade de uma estratégia estruturada de reflexão pós-simulação (Dreifuerst, 2015). O DML favorece aprendizagem ativa ao organizar a reflexão em etapas que explicitam raciocínio, decisões, alternativas possíveis e aplicação do aprendizado. Além disso, o modelo utiliza questionamento de estilo socrático para conduzir o participante à construção de *insights* sobre o próprio desempenho, funcionando também como recurso de modelagem ao processo de descoberta guiada e, portanto, como facilitador do aprendizado do diálogo socrático na formação em TCC.

Instrumentos e materiais

No âmbito dos instrumentos e materiais utilizados neste estudo, foram empregados:

1. Formulário de levantamento sociodemográfico e background.

Elaborado pelos pesquisadores, abrangendo itens como sexo, idade, formação, titulação, ano de conclusão da última titulação, área, local e tempo de atuação no presente momento, experiência docente e tempo de atuação nessa área específica.

2. Escala de avaliação

Duas escalas para avaliação do conteúdo, aplicadas em duas rodadas distintas. Na primeira rodada, utilizou-se uma escala de concordância com opções de resposta “adequado” e “não adequado”, acompanhadas de um espaço para sugestões de melhoria e correções. Na segunda rodada, o instrumento de avaliação foi configurado em formato de escala Likert de 6 pontos, para mensurar o grau de concordância em relação aos itens avaliados, com opções de resposta que variavam de “totalmente inadequado” a “totalmente adequado”. Além disso, em cada questão avaliada, foi incluído um espaço para sugestões e correções, permitindo que os juízes especialistas pudessem oferecer *feedback* qualitativo sobre as respostas consideradas inadequadas.

Procedimento de coleta de dados

No decorrer do procedimento de coleta de dados, duas etapas distintas foram delineadas:

1. Na primeira etapa, dedicou-se à elaboração e planejamento do treinamento, juntamente com a concepção do instrumento de avaliação de conhecimento, por meio de um levantamento bibliográfico. A principal obra de referência adotada foi *Socratic Questioning for Therapists and Counsellors* (Waltman et al., 2017), escolhida por sua relevância na compilação de aspectos essenciais ao desenvolvimento da competência de DS, além de ser recente e redigida por autores conceituados no campo da TCC. O levantamento bibliográfico, que subsidiou a justificativa do treinamento, foi conduzido nas bases de dados PubMed, Scopus, Web of Science e Lilacs, acessadas em fevereiro de 2022, resultando na seleção de três artigos pertinentes (Braun et al., 2015; Kazantzis et al., 2018; Waltman et al., 2017). A escolha desses artigos baseou-se, principalmente, na afinidade entre o conteúdo das pesquisas e os objetivos do treinamento.
2. A segunda etapa consistiu na validação de conteúdo, na qual 05 juízes especialistas da área de psicologia e/ou psiquiatria avaliaram os conteúdos do treinamento e da avaliação de conhecimento. Esse número foi considerado adequado para os propósitos da pesquisa, conforme recomendações estabelecidas (Lynn, 1986). Os convites foram encaminhados

por e-mail, seguidos pelo envio do Termo de Consentimento Livre e Esclarecido (TCLE) aos participantes. A comunicação e o compartilhamento de materiais com os juízes foram realizados por meio da plataforma REDcap®, de propriedade da Faculdade Israelita de Ciências da Saúde (FICSAE). O prazo estipulado para as duas rodadas de avaliações foi de 15 dias corridos após o recebimento do material completo. No entanto, durante a primeira rodada de avaliações, alguns juízes solicitaram uma prorrogação do prazo por mais 15 dias corridos, solicitação acatada pelos pesquisadores e concedida a todos os participantes.

Os critérios de inclusão adotados para a seleção da população de juízes especialistas foram os seguintes: a) formação acadêmica em psicologia e/ou psiquiatria; b) titulação de Doutor em psicologia e/ou áreas afins; c) experiência profissional na abordagem da Terapia Cognitivo-Comportamental (TCC); d) atuação como docente na área de psicologia e/ou áreas relacionadas; e) envolvimento em atividades de pesquisa. No que diz respeito à experiência docente, foram considerados os docentes com pelo menos 5 anos de experiência em TCC.

Os juízes especialistas receberam orientações sobre como proceder e informações sobre os critérios para avaliação dos conteúdos quando tiveram acesso aos materiais para validação. Neste estudo, adotamos os critérios de abrangência, clareza, pertinência, objetividade e praticabilidade, conforme proposto por Alexandre e Coluci (2011).

O procedimento de coleta de dados da segunda etapa foi conduzido por meio da plataforma REDcap® e consistiu em duas rodadas de avaliações. O prazo final estabelecido para a primeira rodada de avaliações foi de 30 dias corridos a partir do envio do material completo. Após esse período, foi realizada uma análise estatística que indicou a necessidade de uma segunda rodada de avaliações, devido à falta de concordância entre os juízes em alguns itens, bem como às sugestões de alterações em partes consideradas inadequadas. Após a análise estatística, os pesquisadores revisaram e justificaram cada uma das sugestões feitas pelos juízes, decidindo quais alterações seriam acatadas ou não. A primeira rodada ocorreu entre os meses de dezembro de 2022 e janeiro de 2023. Na segunda rodada, as avaliações se concentraram nos aspectos que não alcançaram o coeficiente de concordância necessário na etapa anterior, bem como nas alterações sugeridas e aceitas pelos pesquisadores. A partir desse ponto, nenhuma outra alteração foi feita. A segunda rodada de avaliações ocorreu entre os meses de março e abril de 2023.

RESULTADOS

Instrumento de sistematização do treinamento

Foi desenvolvido um documento de sistematização que organiza pedagogicamente o treinamento, descrevendo suas etapas, objetivos, lógica de aplicação e orientações de condução, com a finalidade de favorecer clareza para quem implementa, consistência para quem participa e maior replicabilidade do programa. A organização seguiu o modelo de “5 partes” proposto por Stolovitch e Keeps (2011), utilizado no desenho instrucional por apoiar a aquisição de novos conhecimentos a partir de princípios que orientam o engajamento e a aprendizagem,: explicitação do porquê aprender (razão e expectativa), definição do que será alcançado (objetivos), construção de uma estrutura coerente (sequenciamento e retenção), estímulo à resposta ativa do participante, provisão de *feedback* voltado à tarefa e recompensa (intrínseca e/ou extrínseca) como reforço do envolvimento e da realização.

Com base nessa lógica, o documento foi organizado em: (a) Justificativa, articulando a relevância da competência terapêutica em DS para o processo psicoterápico em TCC e sua relação com objetivos terapêuticos e resultados clínicos (Heiniger; Clark; Egan, 2018; Vittorio et al., 2022; Braun et al., 2015; Kazantzis et al., 2018; Waltman et al., 2017); (b) Objetivos de aprendizagem, apoiados e definidos pela Taxonomia de Bloom (Bloom, 1956) e a partir de sistematizações contemporâneas do DS e de sua aplicação na TCC, com ênfase em competências observáveis ao final do treinamento (Watts et al., 2021) e (c) Etapas do treinamento, planejadas para organizar o percurso formativo de maneira progressiva e coerente, favorecendo compreensão, prática guiada e consolidação do repertório (Stolovitch; Keeps, 2011; Waltman et al., 2021).

Instrumento de aquisição de conhecimento

A “Avaliação do conhecimento do DS em TCC” consiste em um instrumento para avaliar o conhecimento dos participantes. Ele contém uma vinheta clínica que descreve um caso de ansiedade e seis questões relacionadas ao conhecimento e aplicação do DS, sendo quatro objetivas e duas de aplicação.

Cenários de telessimulação

Os cenários foram desenvolvidos considerando o conteúdo a ser treinado e a modalidade de telessimulação com SP, na qual o participante interage com um ator/atriz profissional no papel de cliente, permitindo simular a condução da descoberta guiada a partir da aplicação do DS (Kühne et al., 2018). Como ponto de partida, o facilitador apresenta o contrato de ficção e a netiqueta, com o objetivo de alinhar expectativas, favorecer segurança psicológica e organizar a participação no ambiente on-line, prática descrita e validada por Watts et al. (2021).

O foco central da telessimulação é a aplicação do DS no contexto de descoberta guiada. Para aproximar a experiência simulada da prática clínica, foram incorporados instrumentos típicos da TCC, como o monitoramento de humor (registro emocional) e o registro de pensamentos disfuncionais (RPD), utilizados como suportes para guiar a investigação socrática e a identificação de cognições relevantes durante a sessão.

A vinheta clínica descreve o caso de uma mulher de 35 anos, casada, diretora de uma multinacional, com ansiedade e sintomas associados a ataques de pânico (taquicardia, sudorese, falta de ar e pensamentos catastróficos). O primeiro cenário foi situado na 3ª sessão de psicoterapia, partindo do monitoramento de humor realizado entre a 2ª e a 3ª sessões. Seus objetivos incluíram: demonstrar a condução estruturada do DS, com foco em cognições-chave e compreensão fenomenológica, além de objetivos não técnicos, como usar o monitoramento para orientar a intervenção e sustentar a descoberta guiada, evitando conduções excessivamente diretivas. O perfil do SP foi definido para manter coerência com a apresentação clínica esperada (por exemplo, vestimenta social e sinais comportamentais de ansiedade/hipervigilância).

O segundo cenário foi situado na 4ª sessão, partindo da cognição-chave identificada anteriormente. Nessa etapa, o participante utiliza o RPD para avançar nas etapas do DS (curiosidade colaborativa, síntese e fechamento), evidenciando sua aplicação completa. Entre os objetivos não técnicos, incluíram-se colaboração, escuta empática e comunicação não verbal congruente (expressões faciais suaves). Em ambos os cenários, as interações foram roteirizadas para orientar a atuação do SP e garantir consistência pedagógica, incluindo um “ponto de virada” (mudança planejada durante a simulação) para apoiar o participante a redirecionar o foco ao objetivo educacional central. Cada cenário foi planejado com duração de 20 minutos de interação participante–cliente. A construção e o planejamento dos cenários seguiram diretrizes internacionais para simulação clínica, visando alcançar resultados de aprendizagem e reduzir falhas operacionais (Watts et al., 2021).

Modelo de debriefing

A escolha do modelo de *debriefing* foi baseada em evidências, alinhada aos objetivos de aprendizagem e à modalidade de simulação, contribuindo para padronizar o treinamento do facilitador. Conforme as diretrizes de Watts et al. (2021), o *debriefing* constitui a etapa reflexiva da simulação, realizada imediatamente após a experiência e conduzida por facilitador treinado. Neste programa, o tempo destinado ao *debriefing* correspondeu ao dobro do tempo previsto para cada cenário, favorecendo pensamento reflexivo, *feedback* sobre desempenho e transferência da aprendizagem, em consonância com a teoria experiencial de Kolb (1984).

Considerando tais orientações, adotou-se o modelo *Debriefing for Meaningful Learning* (DML) (Dreifuerst, 2015). A escolha também se justificou por sua abordagem de aprendizagem baseada em questionamento socrático, que acessa o pensamento subjacente às ações e favorece a modelagem da competência em treinamento, promovendo reflexão na ação, sobre a ação e após a ação, além de funcionar como modelo de aplicação do DS ao próprio participante. O DML organiza-se em seis ações (envolver, explorar, explicar, elaborar, avaliar e ampliar) e prevê apoio visual por anotações com marcações por cores: preto (o que ocorreu/foi dito), vermelho (oportunidades de melhoria/erros), verde (pontos positivos/corretos/eficazes) e azul (reflexões,

mudanças e novas perspectivas), além de materiais próprios do modelo. Os pesquisadores obtiveram autorização da autora para uso dos materiais do DML em 16/08/2023.

Validação de conteúdo: proporção de concordância

Quanto à amostra, foi composta por cinco juízes, com idade média de 34 anos (quartil: 31 a 38), sendo dois do sexo feminino (40%) e três do sexo masculino (60%). Quanto ao tempo de atuação clínica em TCC, observou-se mediana de 13 anos (quartil: 10 a 14), com variação de 5 a 14 anos.

As respostas dos especialistas foram analisadas em termos de frequências absolutas e relativas. Cada trecho do treinamento e cada item do instrumento foram avaliados por meio da proporção de adequação, agrupando-se as respostas em “adequado” e “não adequado”. Essa estratégia buscou sintetizar o julgamento dos juízes e orientar revisões nos itens assinalados como não adequados, garantindo maior confiabilidade e validade dos resultados.

Na Tabela 1 pode-se observar a proporção de adequação para as diferentes etapas do treinamento. Itens destacados com percentuais menores referem-se às dimensões de: (i) instrumento de sistematização: clareza e objetividade; (ii) instrumento de aquisição de conhecimento: clareza e pertinência; (iii) cenários: abrangência, clareza, objetividade e pertinência; (iv) netiqueta e contrato de ficção: clareza e pertinência; (v) material de apoio ao cenário: abrangência, clareza e objetividade; (vi) material do facilitador: abrangência, clareza, objetividade e pertinência; (vii) checklist de aderência ao cenário: pertinência; (viii) instrumento de caracterização: clareza; e (ix) *debriefing*: abrangência, pertinência e praticabilidade.

Em relação ao instrumento de aquisição de conhecimento, os itens considerados “não adequados” foram as questões 1, 2 e 5, devido à falta de clareza.

Tabela 1. Proporção de adequação das etapas do treinamento e do instrumento de aquisição de conhecimento

	Abrangência	Clareza	Objetividade	Pertinência	Praticabilidade
Sistematização do treinamento					
Justificativa	85,70%	71,40%	85,70%	85,70%	66,7%*
Objetivos	85,70%	71,40%	85,70%	100,00%	85,70%
Descrição das etapas	85,70%	57,10%	71,40%	85,70%	85,70%
Cenário 1	100%*	42,90%	85,70%	100,00%	42,90%
Perfil do paciente simulado 1	71,40%	57,10%	85,70%	71,40%	85,70%
Frames do cenário 1	85,70%	57,10%	85,70%	100,00%	100,00%
Etapas do cenário 1	85,70%	85,70%	100,00%	85,70%	100,00%
Preparação 1	100,00%	100,00%	100,00%	100,00%	100,00%
Folha de porta 1	100,00%	85,70%	100,00%	100,00%	85,70%
Cenário 2	83,3%*	57,10%	100,00%	100,00%	57,10%
Perfil do paciente simulado 2	100,00%	85,70%	100,00%	100,00%	100,00%
Frames do cenário 2	85,70%	71,40%	100,00%	100,00%	100,00%
Etapas do cenário 2	100,00%	85,70%	100,00%	100,00%	83,3%*
Preparação 2	100,00%	100,00%	100,00%	100,00%	100,00%
Folha de porta 2	100,00%	85,70%	100,00%	100,00%	85,70%
<i>Debriefing</i>	57,10%	85,70%	85,70%	71,40%	71,40%
Plano de ação	100,00%	85,70%	100,00%	100,00%	85,70%

Em negrito: Seis respostas foram apresentadas. Demais casos com sete respostas.

Fonte: Elaboração dos autores.

Tabela 1. Continuação...

	Abrangência	Clareza	Objetividade	Pertinência	Praticabilidade
Avaliação do conhecimento					
Relato de caso	83,3%*	100,00%	100,00%	100,00%	100,00%
Questão 1	100,00%	71,40%	100,00%	100,00%	100,00%
Questão 2	100,00%	71,40%	100,00%	100,00%	100,00%
Questão 3	100,00%	100,00%	100,00%	100,00%	100,00%
Questão 4	100,00%	100,00%	100,00%	100,00%	100,00%
Questão 5	100,00%	42,90%	100,00%	100,00%	100,00%
Questão 6	100,00%	85,70%	85,70%	100,00%	85,70%

Em negrito: Seis respostas foram apresentadas. Demais casos com sete respostas.

Fonte: Elaboração dos autores.

Considerando o Índice de Validade de Conteúdo (IVC) de 0,8 (Altman, 1990), em que coeficientes menores que 0,2 são considerados fracos, de 0,21 a 0,4 razoáveis, de 0,41 a 0,6 moderados, de 0,61 a 0,8 bons e valores superiores a 0,81 muito bons, e adotando-se também a Razão de Validade de Conteúdo (RVC/CVR) para o julgamento de essencialidade dos itens (Lawshe, 1975; Ayre; Scally, 2014), os itens com menor desempenho (Tabela 2) foram revisados conforme as sugestões qualitativas fornecidas pelos juízes e submetidos a uma segunda avaliação.

Observação: Embora a questão 4 tenha recebido concordância na primeira rodada de avaliação, os pesquisadores identificaram uma inconsistência na formulação do item, realizando ajustes redacionais para aumentar sua precisão e coerência com o objetivo do instrumento.

Foram rejeitadas as sugestões de alterações que não estavam alinhadas com o objetivo do treinamento e com o foco do programa, preservando-se a estrutura proposta e as diretrizes metodológicas já estabelecidas para os materiais. Considerando que as respostas (“adequado” e “não adequado”) não foram suficientemente esclarecedoras quanto à adequação de determinados itens, os autores mantiveram a análise qualitativa dos comentários como guia principal para refinar redação, clareza, aplicabilidade e organização dos materiais, garantindo consistência e alinhamento aos objetivos educacionais.

A exemplo, a análise dos comentários evidenciou que as recomendações dos juízes se concentraram, majoritariamente, em ajustes de clareza, detalhamento operacional e padronização, com ênfase na replicabilidade do programa. Entre as sugestões recorrentes, destacaram-se: (i) incluir/antecipar psicoeducação e conceitualização do diálogo socrático (DS) e explicitar o que se entende por condução “estruturada”; (ii) tornar mais explícito o *debriefing* (conceito, momento e dinâmica de execução), incluindo orientações sobre como as perguntas seriam conduzidas e como se organizaria a participação do grupo; (iii) padronizar tempos e descrições entre documentos (especialmente duração de cenário e *debriefing*) e indicar, ao longo das etapas, onde cada material (p.ex., folha de porta) seria utilizado; (iv) aprimorar elementos do desenho instrucional e do cenário para favorecer autenticidade e consistência (p.ex., detalhar demanda e motivo de busca do paciente simulado, revisar a redação de frames e padronizar pessoa/linguagem em instrumentos como o RPD); e (v) refinar itens da avaliação de conhecimento para aumentar a compreensibilidade (p.ex., formatação, marcações e alternativas).

Na segunda rodada, dois juízes não enviaram suas análises, reduzindo o painel para cinco e elevando o critério crítico de concordância para 1,0, conforme (Lynn, 1986), em comparação aos 0,8 da primeira rodada. Nessa etapa, foram reavaliados apenas os itens cujas sugestões de melhoria foram aceitas, sem alterações no instrumento de aquisição de conhecimento. Os itens reavaliados atingiram concordância plena (100%) e índices de CVR compatíveis com a validação do conteúdo do programa.

Na segunda rodada de avaliação, todos os itens reavaliados apresentaram **RVC/CVR = 1,00** nas dimensões de **abrangência, clareza, pertinência, objetividade e praticabilidade**, indicando

Tabela 2. Razão de validade de conteúdo dos itens.

	Abrangência	Clareza	Objetividade	Pertinência	Praticabilidade
Sistematização do treinamento					
Justificativa	0,714	0,429	0,714	0,714	0,333
Objetivos	0,714	0,429	0,714	1,000	0,714
Descrição das etapas	0,714	0,143	0,429	0,714	0,714
Cenário 1	1,000	-0,143	0,714	1,000	-0,143
Perfil do paciente simulado 1	0,429	0,143	0,714	0,429	0,714
Frames do cenário 1	0,714	0,143	0,714	1,000	1,000
Etapas do cenário 1	0,714	0,714	1,000	0,714	1,000
Preparação 1	1,000	1,000	1,000	1,000	1,000
Folha de porta 1	1,000	0,714	1,000	1,000	0,714
Cenário 2	0,667	0,143	1,000	1,000	0,143
Perfil do paciente simulado 2	1,000	0,714	1,000	1,000	1,000
Frames do cenário 2	0,714	0,429	1,000	1,000	1,000
Etapas do cenário 2	1,000	0,714	1,000	1,000	0,667
Preparação 2	1,000	1,000	1,000*	1,000	1,000
Folha de porta 2	1,000	0,714	1,000	1,000	0,714
<i>Debriefing</i>	0,143	0,714	0,714	0,429	0,429
Plano de ação	1,000	0,714	1,000	1,000	0,714
Avaliação do conhecimento					
Relato de caso	0,667	1,000	1,000	1,000	1,000
Questão 1	1,000	0,429	1,000	1,000	1,000
Questão 2	1,000	0,429	1,000	1,000	1,000
Questão 3	1,000	1,000	1,000	1,000	1,000
Questão 4	1,000	1,000	1,000	1,000	1,000
Questão 5	1,000	-0,143	1,000	1,000	1,000
Questão 6	1,000	0,714	0,714	1,000	0,714

*Seis respostas foram apresentadas. Demais casos com sete respostas.

Fonte: Elaboração dos autores.

concordância plena quanto à essencialidade do conteúdo do programa após as revisões implementadas. A concordância entre os especialistas foi avaliada utilizando o coeficiente de concordância AC1 de Gwet, em que as categorias “adequado” e “não adequado” foram analisadas em duas classes. Considerando que houve concordância completa em parte dos itens, o coeficiente atingiu o valor máximo, com limitações para estimação de intervalo de confiança em situações de concordância perfeita. Ainda assim, os resultados permitem inferir que houve concordância adequada entre os especialistas (Gwet, 2014).

DISCUSSÃO

Este estudo desenvolveu e reuniu evidências de validade de conteúdo de um programa de treinamento baseado em telessimulação para o desenvolvimento da competência terapêutica

de diálogo socrático em terapeutas em formação em TCC. A validação por juízes especialistas permitiu identificar pontos de aprimoramento e promover refinamentos ligados sobretudo à clareza, à operacionalização e à aplicabilidade dos materiais, culminando em concordância plena nos itens reavaliados na segunda rodada. Em termos educacionais, o achado central não se limita à aprovação global do conteúdo, mas ao modo como as revisões fortaleceram a coerência instrucional (objetivos–atividades–critérios) e a replicabilidade, atributos necessários quando se pretende implementar intervenções formativas em especializações e educação continuada. Em consonância com Frank, Becker-Haimes e Kendall (2020), treinamentos com maior probabilidade de produzir mudança no desempenho combinam estratégias ativas, prática guiada e *feedback*, enquanto propostas predominantemente expositivas tendem a ser insuficientes para sustentar a transferência para a prática clínica. Nesse sentido, os ajustes orientados pelos juízes aumentam a “prontidão educacional” do programa, ao tornar mais explícitas e observáveis as oportunidades de treino e as devolutivas alinhadas ao desenvolvimento de competências.

Um ponto relevante é que o programa foi estruturado para tornar a telessimulação um dispositivo de ensino, e não apenas uma prática simulada. Conforme Brandão, Collares e Marin, (2014) e Kaneko e Lopes (2019), o ensino por simulação exige equilibrar fidelidade e padronização sem perder o alinhamento aos objetivos educacionais; nesse sentido, contrato de ficção, netiqueta, roteirização do SP e definição de pontos críticos (incluindo “ponto de virada”) funcionam como mecanismos de consistência pedagógica. Também conforme diretrizes internacionais, resultados de aprendizagem dependem de planejamento intencional, alinhamento entre objetivos e desenho do cenário e condições para aprendizagem segura e observável (Watts et al., 2021). Complementarmente, evidências em psicoterapia sugerem que pacientes simulados podem sustentar autenticidade e engajamento quando há preparo adequado do SP e coerência clínica do cenário (Ay-Bryson et al., 2020).

As escolhas relacionadas ao modelo de *debriefing* ampliam o potencial formativo do programa ao favorecer aprendizagem ativa e transferência para situações futuras. Conforme Frank, Becker-Haimes e Kendall (2020), intervenções baseadas em evidências tendem a demandar prática guiada, *feedback* e repetição deliberada para produzir mudança em desempenho; nessa mesma direção, Roth e Pilling (2008) destacam a centralidade de estratégias que tornem o repertório observável e treinável. Assim, o *debriefing* opera como espaço de explicitação de escolhas clínicas, alternativas e critérios, sustentando a aprendizagem a partir do desempenho observado (Dreifuerst, 2015).

Do ponto de vista educacional, o programa contribui em três frentes: (1) oferece um modelo estruturado para ensinar uma competência específica, reduzindo a dependência de formatos exclusivamente expositivos, cuja capacidade de modificar prática clínica é limitada sem treino e *feedback* (Frank; Becker-Haimes; Kendall, 2020); (2) responde à demanda da pós-graduação/educação continuada por estratégias replicáveis e potencialmente escaláveis; e (3) ao incorporar instrumentos típicos da TCC (monitoramento de humor e RPD), aumenta a proximidade com o cotidiano clínico, favorecendo generalização e transferência, objetivo central no desenho de experiências de simulação (Watts et al., 2021; Hayden et al., 2018).

Apesar do potencial, a implementação envolve desafios técnicos e pedagógicos. No plano técnico-operacional, conectividade instável, falhas de áudio/vídeo, latência e desigualdade de familiaridade com plataformas podem comprometer a fluidez do cenário e reduzir presença percebida, exigindo rotinas de checagem, apoio técnico e desenho com tolerância a interrupções (Hayden et al., 2018; Thomas et al., 2021). No plano pedagógico, conforme Watts et al. (2021) e Dreifuerst (2015), a qualidade da facilitação é determinante para resultados educacionais: requer sustentar segurança psicológica, conduzir *prebriefing* consistente, observar marcadores de desempenho e conduzir *debriefing* com precisão, sem converter a reflexão em aula expositiva. Há ainda desafios de fidelidade/padronização: assegurar consistência do SP e coerência emocional/narrativa, sobretudo ao treinar microcompetências como formulação de perguntas, síntese e manejo de respostas; conforme Ay-Bryson et al. (2020), *scripts* claros, treinamento do SP e critérios de variação controlada reduzem ruído e preservam comparabilidade. Por fim, modalidades remotas demandam atenção ética e institucional a privacidade, confidencialidade e armazenamento de dados, especialmente em caso de gravação.

As evidências devem ser interpretadas como iniciais, pois se referem à validade de conteúdo. Essa etapa é necessária no desenvolvimento de materiais educacionais, mas não demonstra, por si só, efetividade para produzir mudança de desempenho nem manutenção ao longo do

tempo. Além disso, houve redução do número de juízes entre rodadas, o que pode limitar diversidade de avaliação e estabilidade dos indicadores; a concordância plena observada é compatível com o critério crítico para painéis menores, mas reduz a variabilidade necessária para estimativas inferenciais mais informativas (Lynn, 1986; Gwet, 2014).

Reconhece-se que a opção por priorizar a validação por especialistas, embora coerente com o objetivo desta etapa, privilegia a consistência interna e a adequação teórica do programa, sem incorporar ainda, de modo sistemático, a perspectiva de aprendizes, facilitadores e atores diretamente envolvidos na experiência educacional e em sua viabilidade operacional. Além disso, a decisão de reavaliar apenas itens com sugestões aceitas, alinhada ao plano de revisão, pode reduzir a oportunidade de testar a robustez do conjunto frente a críticas alternativas. Assim, as conclusões representam um primeiro passo necessário, porém insuficiente, para afirmar impacto educacional, reforçando a necessidade de estudos de pilotagem e implementação.

Pesquisas futuras devem avançar em três direções: (1) estudo piloto com avaliação pré-pós de conhecimento e, principalmente, de habilidades observáveis (por exemplo, rubricas/ checagens de competência aplicadas ao desempenho); (2) dados qualitativos de participantes e facilitadores (aceitabilidade, engajamento, autenticidade e utilidade do *debriefing*); e (3) testes em diferentes arranjos (turmas, instituições e modalidades híbridas) para investigar escalabilidade, custo operacional e consistência de resultados, aspectos relevantes para adoção sustentada de intervenções educacionais (Thomas et al., 2021).

CONCLUSÃO

Ao enfrentar o desafio de transformar uma competência clínica complexa em objetivos, tarefas e critérios observáveis, este estudo reuniu evidências iniciais de validade de conteúdo de um programa de telessimulação para o desenvolvimento do diálogo socrático em TCC. A validação por juízes especialistas não apenas sustentou a consistência teórica dos materiais, como também orientou refinamentos de clareza, operacionalização e aplicabilidade instrucional, resultando em uma versão mais robusta para contextos de pós-graduação e educação continuada.

Para além da validação técnica, os achados reforçam uma implicação educacional relevante: competências terapêuticas frequentemente tratadas como “intangíveis” podem ser operacionalizadas e treinadas de modo sistemático quando são sustentadas por metodologias ativas, prática deliberada e *feedback* estruturado. Nesse sentido, a telessimulação, quando ancorada em desenho pedagógico intencional, pode contribuir para reduzir a dependência de modelos formativos predominantemente expositivos e aproximar o treinamento das exigências reais de desempenho clínico.

Em termos de equidade formativa, o programa apresenta potencial para ampliar o acesso a oportunidades de treino supervisionado em cenários nos quais barreiras geográficas e restrições logísticas limitam experiências presenciais. Ao organizar situações simuladas com SP, critérios de condução e *debriefing* estruturado, a proposta pode favorecer experiências de aprendizagem mais replicáveis e comparáveis, apoiando a formação de terapeutas em diferentes arranjos institucionais e modalidades de oferta (remota, híbrida ou presencial).

Por fim, as evidências devem ser interpretadas com cautela, pois se restringem à validade de conteúdo e incluem limitações relacionadas ao tamanho do painel e à redução de juízes entre rodadas. Esses limites apontam a necessidade de estudos subsequentes, incluindo piloto e investigações de implementação, para examinar efetividade, aceitabilidade e condições de escalabilidade. Ainda assim, a telessimulação se apresenta como uma alternativa metodologicamente estruturada e potencialmente relevante para integrar, de modo crítico e planejado, práticas formativas voltadas ao desenvolvimento de competências terapêuticas em TCC.

REFERÊNCIAS

ALEXANDRE, N. M. C.; COLUCI, M. Z. O. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 16, n. 7, p. 3061-3068, 2011. DOI: <https://doi.org/10.1590/S1413-81232011000800006>. PMid:21808894.

ALTMAN, D. G. **Practical statistics for medical research**. Boca Raton: Chapman and Hall/CRC, 1990. DOI: <https://doi.org/10.1201/9780429258589>.

AY-BRYSON, D. S. *et al.* Can psychotherapy trainees distinguish standardized patients from real patients? A pilot study. **Zeitschrift für Klinische Psychologie und Psychotherapie**, Göttingen, v. 49, n. 3, p. 182-190, 2020. DOI: <https://doi.org/10.1026/1616-3443/a000594>.

AYRE, C.; SCALLY, A. J. Critical values for lawshe's content validity ratio: revisiting the original methods of calculation. **Measurement & Evaluation in Counseling & Development**, Abington, v. 47, n. 1, p. 79-86, 2014. DOI: <https://doi.org/10.1177/0748175613513808>.

BECK, J. S. **Cognitive behavior therapy: basics and beyond**. 3rd ed. New York: The Guilford Press, 2021.

BLOOM, B. S. **Handbook I: cognitive domain**. New York: David McKay Company, 1956.

BRANDÃO, C. S.; COLLARES, C. F.; MARIN, H. F. A simulação realística como ferramenta educacional para estudantes de medicina. **Scientia Medica**, Porto Alegre, v. 24, n. 2, p. 187, 2014. DOI: <https://doi.org/10.15448/1980-6108.2014.2.16189>.

BRAUN, J. D. *et al.* Therapist use of Socratic questioning predicts session-to-session symptom change in cognitive therapy for depression. **Behaviour Research and Therapy**, Oxford, v. 70, p. 32-37, 2015. DOI: <https://doi.org/10.1016/j.brat.2015.05.004>. PMID:25965026.

DREIFUERST, K. T. Getting started with debriefing for meaningful learning. **Clinical Simulation in Nursing**, Oxford, v. 11, n. 5, p. 268-275, 2015. DOI: <https://doi.org/10.1016/j.ecns.2015.01.005>.

FRANK, H. E.; BECKER-HAIMES, E. M.; KENDALL, P. C. Therapist training in evidence-based interventions for mental health: a systematic review of training approaches and outcomes. **Clinical Psychology**, New York, v. 27, n. 3, e12330, 2020. DOI: <https://doi.org/10.1111/cpsp.12330>. PMID:34092941.

GWET, K. L. **Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among raters**. 4th ed. Gaithersburg: Advances Analytics, LLC, 2014.

HAYDEN, E. M. *et al.* Mannequin-based telesimulation: increasing access to simulation-based education. **Academic Emergency Medicine**, Philadelphia, v. 25, n. 2, p. 144-147, 2018. DOI: <https://doi.org/10.1111/acem.13299>. PMID:28846175.

HEINIGER, L. E.; CLARK, G. I.; EGAN, S. J. Perceptions of Socratic and non-Socratic presentation of information in cognitive behaviour therapy. **Journal of Behavior Therapy and Experimental Psychiatry**, Amsterdam, v. 58, p. 106-113, 2018. DOI: <https://doi.org/10.1016/j.jbtep.2017.09.004>. PMID:29055854.

KANEKO, R. M. U.; LOPES, M. H. B. M. Realistic health care simulation scenario: what is relevant for its design? **Revista da Escola de Enfermagem da USP**, São Paulo, v. 53, e03453, 2019. DOI: <https://doi.org/10.1590/s1980-220x2018015703453>. PMID:31166535.

KAZANTZIS, N. *et al.* Socratic dialogue and guided discovery in cognitive behavioral therapy: a modified delphi panel. **International Journal of Cognitive Therapy**, Cham, v. 11, n. 2, p. 140-157, 2018. DOI: <https://doi.org/10.1007/s41811-018-0012-2>.

KOLB, D. A. **Experiential learning: experience as the source of learning and development**. Englewood Cliffs: Prentice-Hall, 1984.

KÜHNE, F. *et al.* Standardized patients in clinical psychology and psychotherapy: a scoping review of barriers and facilitators for implementation. **Academic Psychiatry**, Washington, D.C., v. 42, n. 6, p. 773-781, 2018. DOI: <https://doi.org/10.1007/s40596-018-0886-6>. PMID:29423828.

LAWSHE, C. H. A quantitative approach to content validity. **Personnel Psychology**, New York, v. 28, n. 4, p. 563-575, 1975. DOI: <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>.

LYNN, M. R. Determination and quantification of content validity. **Nursing Research**, New York, v. 35, n. 6, p. 382-385, 1986. DOI: <https://doi.org/10.1097/00006199-198611000-00017>. PMID:3640358.

NAIK, G. *et al.* Impact of use of technology on student learning outcomes: evidence from a large-scale experiment in India. **World Development**, Oxford, v. 127, p. 104736, 2020. DOI: <https://doi.org/10.1016/j.worlddev.2019.104736>.

NEUFELD, C. B. *et al.* Terapia Cognitivo-Comportamental nos cursos de graduação em Psicologia: um mapeamento nacional. **Revista Brasileira de Terapia Comportamental e Cognitiva**, Rio Verde, v. 20, n. 1, p. 86-97, 2018. DOI: <https://doi.org/10.31505/rbtcc.v20i1.1139>.

OVERHOLSER, J. C. Psychotherapy according to the socratic method: integrating ancient philosophy with contemporary cognitive therapy. **Journal of Cognitive Psychotherapy**, New York, v. 24, n. 4, p. 354-363, 2010. DOI: <https://doi.org/10.1891/0889-8391.24.4.354>.

PADESKY, C.A. **Action, dialogue & discovery: reflections on Socratic questioning 25 years later**. Berlin: Padesky, 2019.

ROTH, A. D.; PILLING, S. **A competence framework for the supervision of psychological therapies**. London: University College London, 2008.

RUDÁ, C.; COUTINHO, D.; ALMEIDA FILHO, N. Formação em psicologia: uma análise curricular de cursos de graduação no Brasil. **Revista e-Curriculum**, São Paulo, v. 17, n. 2, p. 419-440, 2019. DOI: <https://doi.org/10.23925/1809-3876.2019v17i2p419-440>.

STOLOVITCH, H. D.; KEEPS, E. J. **Telling ain't training, 2nd edition: updated, expanded, enhanced**. Alexandria: Association for Talent Development, 2011.

THOMAS, A. *et al.* Tips for conducting telesimulation-based medical education. **Cureus**, Palo Alto, v. 13, n. 1, e12479, 2021. PMID:33552792.

VITTORIO, L. N. *et al.* Using socratic questioning to promote cognitive change and achieve depressive symptom reduction: evidence of cognitive change as a mediator. **Behaviour Research and Therapy**, Oxford, v. 150, p. 104035, 2022. DOI: <https://doi.org/10.1016/j.brat.2022.104035>. PMID:35016095.

WALTMAN, S. *et al.* In-session stuck points and pitfalls of community clinicians learning CBT: qualitative investigation. **Cognitive and Behavioral Practice**, New York, v. 24, n. 2, p. 256-267, 2017. DOI: <https://doi.org/10.1016/j.cbpra.2016.04.002>.

WALTMAN, S. H. *et al.* **Socratic questioning for therapists and counselors learn how to think**. New York: Routledge, 2021.

WATTS, P. I. *et al.* Onward and upward: introducing the healthcare simulation standards of best practice™. **Clinical Simulation in Nursing**, Oxford, v. 58, p. 1-4, 2021. DOI: <https://doi.org/10.1016/j.ecns.2021.08.006>.

Contribuições dos autores

RGJ: Conceitualização, Metodologia, Gerenciamento de dados, Coleta de dados, Análise estatística, Análise de dados, Administração do projeto, Preparação visual dos dados, Escrita e revisão. JBB e FALF: Conceitualização, Metodologia, Gerenciamento de dados, Coleta de dados, Análise estatística, Análise de dados, Administração do projeto, Preparação visual dos dados.

Editor: Prof. Dr. José Luís Bizelli

Editora Adjunta Executiva: Profa. Dra. Flavia Maria Uehara