

## RESEARCH ARTICLE

# Rapid transition from in-person to videoconferencing psychotherapy in a counselor training clinic: A safety and feasibility study during the COVID-19 pandemic

Sean B. Hall<sup>1</sup>  | Alise G. Bartley<sup>1</sup> | Julieta Wenk<sup>1</sup> | Annemarie Connor<sup>2</sup>  |  
Suzanne M. Dugger<sup>1</sup>  | Krista Casazza<sup>3</sup> 

<sup>1</sup>Department of Counseling, Florida Gulf Coast University

<sup>2</sup>Department of Rehabilitation Sciences, Florida Gulf Coast University

<sup>3</sup>The Water School, Florida Gulf Coast University

## Correspondence

Sean Hall, 10501 FGCU Boulevard South, Fort Myers, FL 33965, USA.  
Email: sbhall@fgcu.edu

## Abstract

Due to the COVID-19 pandemic, many counselor training clinics rapidly transitioned in-person (IP) services to videoconferencing psychotherapy (VCP). Because VCP is a relatively new technology, more research is needed to establish whether this delivery format is a safe and acceptable substitute for IP services in counselor training clinics. The purpose of this study is to explore questions related to how clients perceive VCP versus IP in terms of credibility and expectancy. Results from this investigation demonstrate that clients who participate in VCP, without first meeting their counselor in person, may initially question the credibility and effectiveness of VCP. However, results demonstrated improvement, in both groups, across the duration of therapy. These findings provide both initial support for the safety of VCP in counselor training clinics and justification for further research.

## KEYWORDS

clinical mental health counseling training center, in person psychotherapy, videoconferencing psychotherapy

## INTRODUCTION

Shortly after the World Health Organization designated COVID-19 as a highly transmissible and potentially deadly illness, experts on infectious disease quickly recognized that the exponential spread of new cases could quickly overwhelm health care agencies (Mallah et al., 2021). In response, the United States healthcare system embraced telehealth technology to ensure access to care while minimizing the likelihood of community transmission (Gruber et al., 2021). Similarly, colleges and universities also began transitioning online. In this context, university-based counselor training clinics (UBCTC) were forced to either suspend clinical operations indefinitely or abruptly transition in-person (IP) services to a synchronous, videoconferencing format (Hames et al., 2020). Under normal circumstances, training clinic directors would carefully weigh multiple factors (e.g., legal and ethical issues, clinical care considerations, training quality, and procedural

amendments) when deciding whether to make this switch. In the wake of the pandemic, however, critical discussions about the safety and effectiveness of videoconferencing psychotherapy (VCP) were temporarily suspended as the risk of COVID-19 infection, associated long-haul symptoms, and possibility of death outweighed the potential risks.

Due to the COVID-19 pandemic, the use of VCP in training clinics has accelerated (Dyason et al., 2019). Because little is known about whether the relative inexperience of novice counselors impacts how VCP is implemented, researchers must critically examine the effectiveness and feasibility of VCP in this unique context.

## Effectiveness of VCP

Although not specific to UBCTCs, emerging evidence suggests that VCP may be an effective extension of IP

delivery. For example, systematic reviews of randomized, quasi-experimental, and uncontrolled clinical trials indicate that VCP may be associated with medium to large effects on depressive symptoms (Berryhill, Culmer, et al., 2019) and small to large effects on symptoms of anxiety (Berryhill, Halli-Tierney, et al., 2019). Similarly, a meta-analysis of 16 studies evaluating the effectiveness of cognitive behavioral VCP further demonstrated that, relative to various control conditions, VCP appears to produce moderate effects across multiple outcome measures and presenting concerns (Matsumoto et al., 2021). Finally, data from a recent multilevel meta-analysis additionally reported no significant differences in outcomes between clients receiving either IP or VCP services (Batastini et al., 2021).

Although current findings appear to suggest that VCP is effective, additional research is needed to increase the strength and methodological diversity of evidence. For example, a majority of existing studies were formulated as superiority trials. As such, one can only claim that VCP is unlikely to be more effective than IP services. Questions about whether the effects produced by VCP are essentially the same or, at least, not meaningfully worse than IP are better answered by equivalence and noninferiority designs (Wellek, 2010). Although the number of equivalence or noninferiority trials are growing (see Norwood et al., 2018), more high-quality studies are needed before any conclusions can be made about whether VCP is a complementary or inferior substitution for IP delivery.

## Safety considerations

Ethically, counselors are obligated to promote well-being and mental health while simultaneously preventing harm (American Counseling Association, 2014). As such, counselors must consider issues of safety when selecting interventions. Safety is a state where the probability of unnecessary harm is reduced to a minimally acceptable level. The degree of safety is derived from a risk assessment, which involves estimating the likelihood of negative effects associated with a given intervention (Runciman et al., 2009). Although practitioners and supervisors have long sought to minimize potential harm (Barlow, 2010), this topic, until recently, has been mostly neglected in the professional literature (Lilienfeld, 2007). Although emerging research is beginning to enrich our understanding of some adverse events and harmful effects associated with psychotherapy, little is currently known about the safety of VCP.

## Acceptability

Feasibility designs assess whether and perhaps how a larger future trial should be undertaken and are often designed to evaluate a wide range of questions related to acceptability, recruitment, data collection, analysis, and operational procedures (Eldridge et al., 2016). Studies of acceptability, in

particular, evaluate the degree to which clients and clinicians perceive a treatment to be suitable based upon their personal reactions to it. Of the seven proposed components of acceptability, two are relevant to the present study. First, *perceived effectiveness* represents the subjective likelihood that VCP will achieve its intended purpose, and *opportunity costs* represent the benefits, profits, or values one must forfeit in order to engage in VCP. These components are operationalized in various ways across the literature (Sekhon et al., 2017).

For this investigation, perceived effectiveness was operationalized as treatment credibility (TC) and outcome expectancy (OE). TC refers to how the client perceives the legitimacy, utility, or adequacy of a given intervention (Constantino et al., 2018), while OE is the degree of benefit one anticipates they will receive from participating in the treatment (Constantino et al., 2018). Under a contextual model of psychotherapy, both TC and OE are common factors. TC is enabled during early interactions between the client and counselor, where the provider and setting characteristics convey an overall sense of trustworthiness through various cues such as active listening, empathic responding, professionalism of staff, comforting atmosphere, and degrees/credentials displayed on the wall. Moreover, all forms of psychotherapy are thought to enable the expectation that relief will result from participation and active engagement in the counseling process (Wampold & Imel, 2015).

This study operationalized opportunity costs using premature termination (PT) and treatment refusal (TR). Conceptually, PT is thought to occur when the perceived costs of treatment exceed its expected benefit. For this investigation, PT is formally defined as the unilateral withdrawal from counseling before achieving reliable and/or clinically significant change (Hatchett & Park, 2003; Swift et al., 2009). Last, TR represents the failure to initiate services after scheduling the first appointment. A high rate of TR may reflect a mismatch between client preferences and the available service or the presence of barriers including inconvenience, cost, or negative outcome expectations (Swift et al., 2017). Currently, the delivery of VCP is still in the early phases. Very little is known about acceptability within this medium. Even less is known about how each of these manifest when VCP is delivered in the context of a training clinic setting.

## Implementing evidence-based practices in UBCTCs

UBCTCs offer specialized counseling services to the community and promote the professional development of trainees through experiential learning. Within this setting, a majority of services are delivered by novice graduate students working under the supervision of experienced professionals. UBCTCs are unique from other treatment contexts in their effort to balance parallel, and sometimes competing, priorities. On the one hand, UBCTCs are obligated to deliver competent and

ethical services, monitor effectiveness, and minimize harm (ACA, 2014). On the other, they are tasked with creating a rich learning environment for inexperienced trainees that promotes the gradual integration of formal and experiential knowledge (Norman et al., 2018).

Emerging evidence suggests that the effectiveness of an intervention depends, in part, upon the context in which it is delivered. For example, a number of organizational factors such as neighborhood (Firth et al., 2020), supervision (Gerstenblith et al., 2021), trainee (Erekson et al., 2017), and/or clinic characteristics (Delgadillo et al., 2016) appear to influence the implementation of counseling and its associated impacts. Because UBCTCs are characterized by unique principles distinguishing them from other contexts, the extent to which empirically derived inferences generalize to and from this setting remain uncertain (Dyason et al., 2019). For example, when compared with other clinical settings, clients who receive services in training clinics appear to demonstrate smaller gains and a slower rate of recovery (Callahan & Watkins, 2018). Moreover, significant differences in the rate of PT have also been observed across treatment settings (Swift & Greenberg, 2012) with training clinics consistently reporting among some of the highest rates relative to others (Al-Jabari et al., 2019; Callahan et al., 2014). Such findings make it difficult to anticipate how contextual variables may influence the implementation of VCP in this context, and how these effects might influence both the therapeutic alliance and clinical outcomes.

## Purpose

The COVID-19 pandemic has accelerated the adoption of VCP in training clinics. While evidence supporting the effectiveness of VCP is becoming stronger, there remains a lack of research outlining the safety and acceptability of VCP when delivered by novice counselors. This limitation is notable as the implementation of psychotherapy within a UBCTC may be moderated by multiple factors including trainee inexperience, self-efficacy, supervision, and unique setting characteristics (e.g., constrained hours of operation, more stringent inclusionary/exclusionary criteria, etc.). As a result, inferences derived from research conducted in other contexts may not be transferable to this setting.

Based on the present gaps in our knowledge, a comprehensive evaluation of VCP is warranted. However, given the early stage of research, empirical efforts must first focus on basic questions related to acceptability and safety. Such steps are necessary for refining the clinical and administrative procedures needed for designing and implementing large scale clinical testing (Gallo et al., 2013). The following research questions were developed to guide our design and analysis:

Research question 1. Do clients perceive VCP as a credible alternative to IP services?

Research question 2. Do clients expect that VCP and IP services achieve equivalent therapeutic outcomes?

Research question 3. Do clients who receive VCP change at a comparable rate and direction to those who receive IP services?

## METHOD

### Procedures

After receiving institutional review board approval, data were collected from participants who received services in an outpatient counselor training clinic between 2019 and 2020. The counseling clinic is operated by the Department of Counseling housed within a mid-size southeastern university and accredited by the Council for the Accreditation of Counseling and Related Programs (CACREP). To ensure the confidentiality of participants, data were anonymized by clinic leadership before being made available to the research team for analysis.

For the present study, all sessions were facilitated by 11 graduate students who had progressed into the internship phase of their training while working toward a master's degree in clinical mental health counseling. All interns received weekly supervision from two licensed mental health counselors (author two and three, respectively). Prior to the initiation of services, clients participated in an initial level of care screen. Individuals presenting with severe and persistent mental illness, active suicidality, or substance use disorder(s) received immediate referrals to external providers.

To facilitate comparisons, participants were identified as belonging to one of two groups. The IP group comprised of clients who participated in at least two face-to-face meetings with the counselor prior to a state mandated safer-at-home order. Once the mandate was applied, clients in the IP group were invited to immediately continue receiving services remotely. In contrast, the VCP group consisted of clients who had enrolled in counseling services after initiation of the safer-at-home order and had never met with their assigned provider in-person. More simply, the IP group started face-to-face before transitioning online, while the VCP group exclusively received counseling through an online delivery format. All sessions were 45–50 min in length.

In this setting, cases were assigned based on the identified presenting problem, diagnosis, intern experience/knowledge, and clinical complexity. Initial sessions focused on establishing the therapeutic alliance and gathering clinical data for deriving a case formulation. The clinic operates within a contextual model of psychotherapy as opposed to a biomedical model (Wampold & Imel, 2015). Under this framework, change is presumably enabled via the interaction among three pathways (1) the relationship, (2) specific ingredients, and (3) expectations. The contextual model contends that the underlying mechanism of change in therapy is less related to therapeutic specificity, but more to those processes common across all modalities that invite clients to engage in new, health promoting behaviors (Finsrud et al., 2021). The contextual model is synthesized from the empirical process and outcome research. Although preliminary evidence supports

its essential tenets it is continuously updated to reflect new research (Wampold, 2015; Norcross & Lambert, 2018).

The Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) was administered at the beginning of all sessions, and the Credibility/Expectancy Questionnaire (CEQ; Devilly & Borkovec, 2000) was administered at the beginning of the first VCP session.

## Measures

### Modified CEQ

The CEQ measures both the degree to which a treatment is viewed as logical and effective as well as the amount of improvement clients believe will result from participating in the intervention (Devilly & Borkovec, 2000). The original CEQ is a six-item self-report measure relying on a unipolar 9-point Likert-type rating scale (with labeled midpoint and endpoints) and a unipolar 10-point percentage rating scale (scale ranges from 0 to 100% with all options labeled).

In light of previous research indicating a larger number of response categories may increase cognitive load, resulting in both a lower completion rate and difficulty understanding and distinguishing between the various response categories (Schaeffer & Presser, 2003), the clinical team modified the response categories for all items measured with a Likert-type rating scale (i.e., items 1, 2, 3, and 5). Specifically, the number of response categories for the Likert-type rating scale were collapsed from 9 to 5, while retaining the original mid- and endpoint labels. The impact of this modification is unclear, but perhaps trivial, as recent evidence suggests amendments to the response scale tend to be most problematic for measures with verbal labels attached to all response categories (Steinberg & Rogers, 2020). The original 10-point percentage rating scale for items 4 and 6 was retained. Cronbach's  $\alpha$  was used to estimate internal consistency for the modified credibility,  $\hat{\alpha}_M = 0.80$ , 95% CI (0.65, 0.89), and expectancy subscales,  $\hat{\alpha}_M = 0.79$ , 95% CI (0.61, 0.87). For reference, results from the original multiple-study validation article reported  $\alpha$  estimates of 0.81 and 0.86 for the credibility factor as well as 0.79 and 0.90 for the expectancy factor (Devilly & Borkovec, 2000). For this investigation, a composite of standard scores was constructed for each subscale prior to analysis.

### Patient health questionnaire

The nine-item PHQ-9 assesses the frequency of depressive symptoms over the past two weeks along a 4-point Likert type scale. Higher scores indicate the presence of more frequent depressive symptoms. Depression severity is most frequently measured as the sum score for all items. Recent meta-analytic data indicate a clinical cutoff of at least 10 to maximize diagnostic accuracy (e.g., sensitivity between 0.83 and 0.92 and specificity between 0.82 to 0.88; Levis et al.,

2019). Empirical data suggest the PHQ-9 is suitable for longitudinal analysis, as the factorial structure is likely to remain invariant across sex, race/ethnicity, education level, and time (Guo et al., 2017; Patel et al., 2019; Richardson & Richards, 2008). Prior to hypothesis testing, raw scores on the PHQ-9 at each time point were mapped onto a common depression metric developed as part of a National Institutes of Health (NIH) initiative to ensure comparability across measures of depression frequently used studies of health outcome (Choi et al., 2014). In the current study, internal consistency estimates of Cronbach's  $\alpha$  across each measurement point ranged from 0.80 to 0.90.

### Trajectory of change

A deteriorating trajectory was defined in relation to the naturalistic course of the primary clinical target. Although symptoms of depression are likely characterized by heterogeneous growth patterns, a majority of clients with mild to moderate symptoms appear to follow a stable or improving trajectory (Essau et al., 2020; Musliner et al., 2016). Therefore, an increase in severity ratings on the PHQ-9 represents deterioration, while reduced severity ratings reflect improvement. A flat pattern of change can reflect either the capacity of an intervention to protect against the worsening of symptoms or no change attributable to the intervention (Dimidjian & Hollon, 2010). For this investigation, multiple indicators were used to explore whether clients in the VCP group experience a trajectory of change consistent with those seen among clients in the IP group.

The rate of improvement and deterioration were defined as the proportion of clients in the sample who demonstrated reliable decreases or increases in severity ratings on the PHQ-9 respectively. Reliable change is the degree to which fluctuations in the outcome measure are attributable to real change in the latent variable as opposed to measurement error (Jacobson & Truax, 1991). Improvement was defined by a minimum 5-point decrease in severity ratings on the PHQ-9 while deterioration reflects the opposite. Recovery was defined as crossing the threshold between clinical and nonclinical severity scores and operationalized using the reliable and clinically significant change (RCSC) index. Based on previous research (Levis et al., 2019; McMillan et al., 2010), RCSC was defined as a minimum pretreatment score of 10, a maximum posttreatment score of 9, and an improvement of at least 5 scale points. Finally, both groups were evaluated using a longitudinal growth model where the number of attended sessions were nested within clients, who were nested within the treatment delivery method.

### Premature termination and treatment refusal

A high rate of PT may indicate the presence of a potentially harmful intervention (Dimidjian & Hollon, 2010) or may indicate a lack of acceptability (Sekhon et al., 2017). For this

investigation, PT was operationalized as failure to achieve clinically significant change (CSC) by the last recorded score on the PHQ-9. Last, TR was defined as failure to initiate services after scheduling the first appointment.

### Statistical software

Data and descriptive statistics were collected and computed in Microsoft Excel (version 1808) and inferential analyses and figures were generated in R (version 4.1.1).

## RESULTS

### Sample and setting characteristics

Upon initiation of the state COVID-19 lockdown, clinic leadership immediately began transitioning all services to a remote delivery format. At this time, 100 clients who had been receiving services in-person were invited to immediately switch over to a VCP delivery format. Of these, 52 agreed to switch, 41 specifically refused VCP, and seven declined to continue in services for reasons unrelated to VCP. Of those who agreed, 20 of the 52 clients were excluded from the study due to either not meeting the minimum age criterion (>10 years of age) or due to incomplete clinical documentation. The 32 remaining clients were categorized into the IP group based on the criteria outlined above. After initiation of the state lockdown the clinic enrolled 21 new clients who were then categorized into the VCP group. Table 1 provides demographic characteristics of the initial participant pool.

### Perceived credibility of VCP services

Preliminary analyses suggested that credibility ratings in both groups were unlikely to follow a normal distribution. In response, a quantile comparison (QC) test was performed. Formally, QC tests belong to the robust family of statistics developed as an alternative set of methods for comparing location estimates when assumptions for t-tests, analysis of variance, and/or Mann–Whitney–Wilcoxon tests are unmet (Mair & Wilcox, 2020). QC tests have two advantages. First, the QC testing procedure yields quantile estimates spanning the entire range of the response distribution. This allows for investigators to examine points along the distribution where response densities between both groups may either converge or deviate from one another. Moreover, prior research also demonstrates that quantile estimates tend to be robust, particularly in the context of small samples ( $n \geq 20$ ), violations of normality, and ties in the data (Wilcox et al., 2014). For the present study, Harrell-Davis quintile estimates were computed for each group. The distances between group location estimates were then evaluated using the percentile bootstrap method (for details of this procedure see Wilcox

TABLE 1 Demographic characteristics for clients in the IP ( $n = 32$ ) and VCP ( $m = 21$ ) groups

Demographic characteristics	Treatment group			
	IP		VCP	
	<i>n</i>	%	<i>m</i>	%
Ethnicity				
Caucasian/White	24	75%	9	42.9%
Hispanic/Latino	4	12.5%	9	42.9%
African American	1	3.1%	–	–
Asian American/Pacific Islander	1	3.1%	–	–
Caribbean Islander	1	3.1%	–	–
Middle Eastern	1	3.1%	–	–
Biracial	–	–	2	9.5%
No response	–	–	1	4.8%
Sex				
Female	26	81.2%	15	71.4%
Male	6	18.8%	6	26.6%
Education level				
Bachelor degree	11	34.4%	10	47.6%
High school	10	31.2%		
Less than high school	8	25%	5	23.8%
Associate degree	3	9.4%	6	28.6%
Age				
	M (SD)		M (SD)	
	28 (15.4)		29 (11.6)	

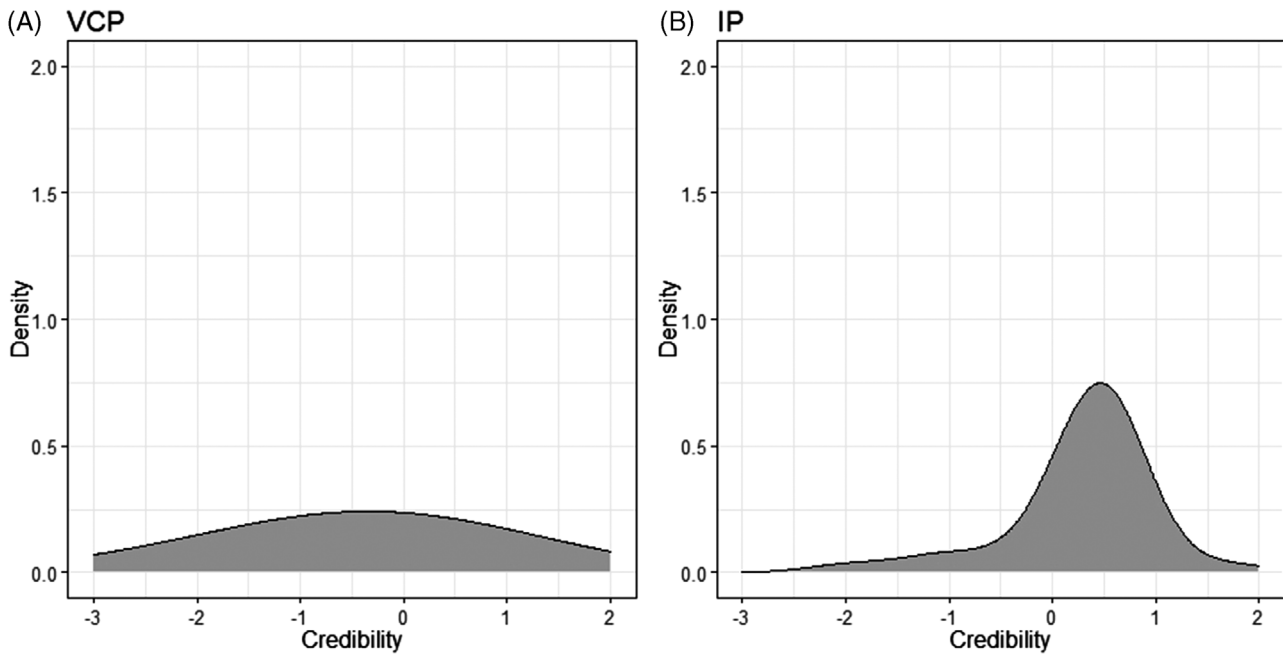
TABLE 2 Parameter estimates for quantile comparisons of credibility scores between clients receiving IP ( $n = 29$ ) or VCP ( $n = 20$ ) services

Quantile	VCP	IP	VCP–IP	95% CI	<i>p</i>
0.10	–1.67	–0.73	–0.94	–2.18	0.11
0.25	–1.08	0.16	–1.23	–2.07	0.002*
0.50	–0.40	0.53	–0.93	–1.52	0.004*
0.75	0.46	0.55	–0.09	–0.85	0.012*
0.90	0.55	0.64	–0.09	–0.91	0.008*

Note: *p* value is less than the critical value; computation discrepancies due to rounding.

et al., 2014). Because group location estimates, within each quantile, are tested simultaneously, the *p* value is adjusted to exert strong control over the familywise error rate (Hochberg, 1988). Due to missing observations on the CEQ, the sample consisted of 29 subjects in the IP group and 20 in the VCP group.

Results are listed in Table 2. These findings demonstrate the IP group had a greater concentration of subjects responding with higher levels of perceived credibility relative to the VCP group. To visualize these differences, Figure 1 depicts the estimated density curves for client perceptions of TC across both groups.



**FIGURE 1** Estimated distribution of credibility scores. Panels (A) and (B) plot the estimated density of credibility scores provided by clients receiving either VCP or IP services respectively. Relative to response patterns observed in the VCP group, clients in the IP group more frequently endorsed higher levels of perceived credibility.

### Research question 2: expectancy of improvement

A Welch's *t*-test was conducted to test differences in expectancy between treatment groups. We compared mean differences in the expectancy of symptom improvement between clients in either the VCP ( $M = -0.5$ ,  $SD = 1.1$ ) or the IP ( $M = 0.3$ ,  $SD = 0.5$ ) group. Results indicated the difference in mean scores on the expectancy subscale was statistically significant,  $t(25.13) = -2.77$ ,  $p = 0.008$ , with a large effect-size, Hedge's  $g = -0.83$ , 95% CI  $[-1.52, -0.31]$ . In summary, clients who did not previously meet their counselor in-person prior to the initiation of VCP expected less benefit from counseling than clients who had.

### Research question 3: trajectory of change

In the present study, 24% (11 out of 46) of clients ( $IP = 4$  and  $VCP = 7$ ) presented with clinically elevated scores on the PHQ-9. Moreover, approximately 17% (eight out of 46) of the sample demonstrated reliable improvement ( $IP = 6$  and  $VCP = 2$ ) and 11% (five out of 46) achieved both reliable and clinically significant improvement ( $IP = 3$  and  $VCP = 2$ ). Also, an estimated 4% (two out of 46) of clients reported a reliable worsening of PHQ-9 severity ratings. Both of which were in the IP group. The rate of PT across the entire sample was 8% and the modal number of completed sessions was seven ( $Min = 1$  and  $Max = 24$ ). The top three reasons for discharge were completion of treatment goals (41.2%), referred to another provider in the

clinic (29.4%), and referred to higher level of specialized care (5.9%).

Next, we used a repeated-measures, time-series design to compare the trajectories of change between VCP and IP delivery. Because data were collected in a clinical setting, clients differed in the number, frequency, and intervals between sessions. As such, we fitted a two-level longitudinal growth model to the sample data where the number of counseling sessions ( $Mode = 7$ ,  $Min = 1$ ,  $Max = 24$ ) were nested within clients. Missing data for all administrations of the PHQ-9 were observed for one subject in each group leaving a sample of 51 ( $IP = 31$  and  $VCP = 20$ ). Group membership (coded as  $VCP = 0$  and  $IP = 1$ ) was specified as a fixed effect. The time predictor, session, was anchored at 0 and model parameters were computed using maximum likelihood estimation. Fixed effects were tested using Student's *t*-distribution with degrees of freedom approximated using the Kenward-Rogers method. Nested model comparisons were tested for improvement in model-data fit with the likelihood ratio test, which evaluates change in the  $-2$  Log Likelihood along a chi-square distribution with degrees of freedom equal to the number of additional parameters included in the more complex model. Results from the PHQ-9, at each time point, are expressed as standard scores.

### Preliminary analyses

To investigate the direction of change over time, model 1 was initially fitted to the data (see Table 3 for indices of model fit and related parameter estimates). First, the influence of fixed

TABLE 3 Model comparisons and parameter estimates for the analysis of longitudinal growth

Fixed effects	Model 1			Model 2			Model 3		
	Coefficient	95% CI	<i>p</i>	Coefficient	95% CI	<i>p</i>	Coefficient	95% CI	<i>p</i>
Level 1									
Intercept, $\beta_{00}$	0.62	(0.32,0.92)	<0.001	0.51	(0.12, 0.91)	0.01	0.68	(0.25, 1.11)	0.003
Tx group, $\beta_{01}$				0.22	(-0.26, 0.71)	0.38	-0.14	(-0.74,0.46)	0.64
Level 2									
Session, $\beta_{10}$	-0.07	(-0.10, -0.04)	<0.001	-0.08	(-0.11, -0.04)	<0.001	-0.13	(-0.19, -0.07)	<0.001
Session $\times$ Tx, $\beta_{11}$							0.08	(0.002, 0.15)	0.04
Random effects									
	SD	Variance Component		SD	Variance Component		SD	Variance Component	
Level 1, $\epsilon_{ij}$	0.5	0.21		0.5	0.20		0.5	0.20	
Initial status, $r_{0j}$	0.9	0.77		0.9	0.80		0.9	0.77	
Growth rate, $r_{1j}$	0.1	0.00		0.1	0.01		0.1	0.00	
Correlation ( $\pi_{0j}, \pi_{1j}$ )									
	$\rho$			$\rho$			$\rho$		
	-0.62			-0.64			-0.60		
Model fit									
	Deviance	Parameters	AIC	Deviance	Parameters	AIC	Deviance	Parameters	AIC
	470.4	6	482.4	469.7	8	483.7	465.7	8	481.7

Confidence intervals estimated using parametric bootstrap with 1000 resamples. Deviance = -2LL; AIC = Akaike Information Criterion.

effects was examined. On average, overall participant scores on the PHQ-9 at the initial measurement period ( $\beta_{00} = 0.62$  95% CI [0.32,0.92]) corresponded to mild depression severity (raw PHQ-9 scores range from 5 to 9). A significant downward trend in depression severity was observed across both groups ( $\beta_{10} = -0.07$  [-0.10, -0.04],  $t(25) = -4.27$ ,  $p < 0.001$ ).

Given the influence of client, counselor, and process-related factors on overall outcome, intercept and slope parameters were expected to vary across participants. To test this assumption, model 1 was compared first to a random intercept model and then to a random slope model. If model 1 demonstrated superior fit, improvement is attributed to either the random slope coefficient or the random intercept respectively.<sup>1</sup> Results demonstrated improvements in fit when comparing model 1 to both the random intercept ( $\chi^2(2) = 15.512$ ,  $p < 0.001$ ) and random slope models ( $\chi^2(2) = 115.58$ ,  $p < 0.001$ ). These findings indicate significant variation in both the initial status of depression severity ( $r_{0j} = 0.88$  [0.62, 1.08]) and in the overall rate of change across the episode of care ( $r_{1j} = 0.07$  [0.03, 0.10]).

### Main analysis

To compare differences in the initial status and rate of change between treatment groups, model 2 was formulated with treatment group added to the intercept.<sup>2</sup> For this analysis,

<sup>1</sup> For space considerations neither model is tabled. For this analysis, the random intercept model was specified as  $\text{lmer}(\text{PHQ9} \sim \text{Session} + (1 | \text{Client}), \text{data} = \text{df})$ . The random slope model was specified as  $\text{lmer}(\text{PHQ9} \sim \text{Session} + (0 + \text{Session} | \text{Client}), \text{data} = \text{df})$ .  
<sup>2</sup> Model 2 was specified as  $\text{lmer}(\text{PHQ9} \sim \text{Session} + \text{Tx} + (1 + \text{Session} | \text{Client}), \text{data} = \text{df})$

VCP was specified as the referent category. A comparison between model 1 and model 2 revealed no improvement to model fit ( $\chi^2(1) = 0.54$ ,  $p = 0.387$ ). Next, model 3 was expressed with a cross-level interaction between the time variable and treatment group.<sup>3</sup> Formally, model 3 was defined as

$$L1 : Y_{ij} = \pi_{0j} + \pi_{1j} \text{Session} + \epsilon_{ij}$$

$$L2 : \pi_{0j} = \beta_{00} + \beta_{01} \text{Tx}_j + r_{0j}$$

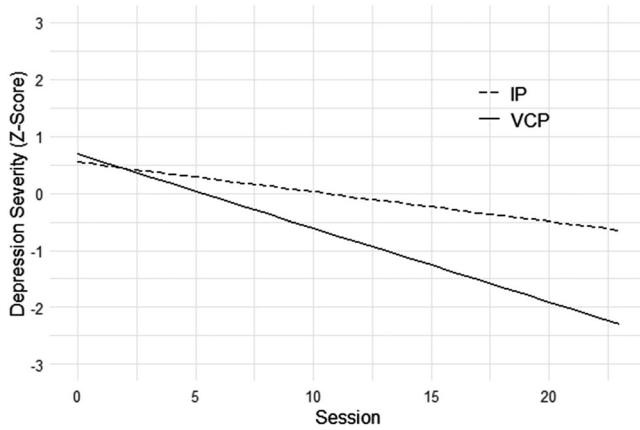
$$\pi_{1j} = \beta_{10} + \beta_{11} \text{Tx}_j + r_{1j}$$

We then compared models 2 and 3 to evaluate the effect of group membership on the rate of change. Results indicated that the addition of the cross-level interaction was significant ( $\chi^2(1) = 4.35$ ,  $p = 0.045$ ). Taken together, these findings suggest that initial severity ratings on the PHQ-9 did not differ across groups, but did appear to differ in the rate of improvement. Figure 2 plots the predicted rate of change in depression severity across treatment groups.

### DISCUSSION

Early in the COVID-19 pandemic, training clinic directors sought to ensure continued access to care while simultaneously mitigating the risk of disease transmission. Within this context, VCP presented a technological solution capable of resolving the competing priorities balanced by clinic directors, including concerns for client welfare, counselor safety,

<sup>3</sup> Model 3 was specified as  $\text{lmer}(\text{PHQ9} \sim \text{Session} * \text{Tx} + (\text{Session} | \text{Client}), \text{data} = \text{df})$



**FIGURE 2** Predicted values of depression severity over time by treatment group

trainee development, and community outreach. However, as the world adapts to the continued threat of COVID-19, clinic directors must reassess the risk–benefit balance. Because VCP is a relatively new technology, more research is needed to establish whether this delivery format is a safe, acceptable, and ultimately effective substitute for IP services in the training clinic environment.

Results from the current study suggest approximately 41% of clients enrolled in IP services prior to the state mandated safer-at-home order specifically declined to continue services in VCP. The elevated refusal rate in our sample may suggest a preference for IP services. Formally, preferences refer to activities, treatments, or therapist qualities that clients want or hope to experience during therapy (Swift et al., 2018). These findings may suggest that some clients believe VCP, when delivered in this context, is unable to provide the experience they want or hope to have. It can further be inferred that the observed sample consisted largely of clients who either had a favorable preference toward VCP or no preference toward either delivery format.

Results further suggest that clients who began VCP, without first meeting their counselor in-person, were less likely to view VCP as a suitable alternative. Clients in this group were also less likely to expect this delivery format would ameliorate their presenting concern(s). Similar skepticism was not observed among clients who began therapy in-person before making the transition to remote services. Perceived TC requires knowledge, observation, and/or direct experience with the clinic, counselor, and/or treatment. Therefore, authors sometimes distinguish between principle (credibility of a treatment in principle), initial (before first session), and/or emergent (after initial session) subtypes (Hardy et al., 1995; Constantine, Coyne, et al., 2018). Outcome expectation (OE) is a separate, but related construct. Formally, OE does not rely on prior knowledge and may exist before any interaction with a treatment or provider (Constantino et al., 2018). As such, these findings may suggest that credibility established during IP therapy may moderate initial skepticism toward VCP during the switch between delivery methods.

These findings may also suggest that even among clients with favorable or neutral preferences toward VCP, some may begin services despite being skeptical of its effectiveness.

Last, although the majority of clients in both groups reported either an improving or flat trajectory of change on the PHQ-9, approximately 4% ( $n = 2$ ) of clients, both in the IP group, appeared to deteriorate. Moreover, based on the rate of PT, approximately 8% of clients receiving services in this context may perceive that the costs associated with VCP outweigh its anticipated benefits and ultimately choose to withdraw from services before achieving clinically significant change. Taken together, these findings may suggest that, for clients who do not hold strong preferences against VCP, this delivery format may be administered safely and is perceived by many to be an acceptable alternative. However, more research is needed to better understand why some who receive VCP in this setting appear to get worse, why some do not appear to follow an improving trajectory, and why others perceive little benefit in continued participation.

## Implications

It is well established that strong preferences, in favor of a specific treatment, may cause recruitment challenges for clinical trials if a large proportion of clients refuse randomization (Howard & Thornicroft, 2006). The observed refusal rate may therefore have implications for the implementation of future research if participants randomized to the VCP group are at a higher risk of TR. For instance, recent reviews suggest matching clients with their preferred treatment is likely to enhance the therapeutic alliance (Windle et al., 2020), increase adherence, and to exert small, but robust, improvements in therapeutic outcome. It is also known that the exclusion of clients with strong preferences may yield biased measures of target outcomes and limit the degree to which the effects of a given treatment can be generalized beyond the individual trial (Swift et al., 2018).

The preference effect can be estimated in several ways. Perhaps the most direct measure is obtained through the match/no-match design. Under this strategy, researchers compare outcomes for clients who were matched to their preferred therapy against those who were not (Manthei et al., 1982). Another protocol is the partial randomized preference trial (PRPT; Brewin & Bradley, 1989). In this design, clients who refuse randomization (clients with a preference) are assigned to their preferred treatment and compared with clients randomized as usual (clients without a preference; Swift & Greenberg, 2015). Recent meta-analytic findings suggest both strategies produce lower preference effects than clients randomized to treatment groups in the usual way (Swift et al., 2018). Based on findings from the present study, two recommendations can be made. First, additional research using both qualitative and match/no-match designs is needed to better understand the nature and magnitude of the preference effect on both the therapeutic alliance and clinical outcome.



Second, investigators are encouraged to use PRPT designs before pursuing traditional randomized group allocation procedures.

Various operational definitions have been used to measure PT. These include duration-based, failure to complete, the missed appointment method, and therapist determination. Notably, the degree to which each of these methods are tapping a common latent construct remains uncertain (Swift & Greenberg, 2015). For example, therapist determination relies exclusively on therapist judgment, which may be attenuated (AEgisdottir et al., 2006) by implicit cognitive biases (Bowes et al., 2020) that tend to persist regardless of how much practical experience a therapist accumulates during their career (Spengler & Pilipis, 2015). As such, when relying solely on clinical judgment, therapists may struggle to recognize client improvement or deterioration (Hatfield et al., 2010). This may result in disagreements between the client and counselor about how much change is needed before counseling can reasonably be discontinued (Swift & Callahan, 2008). Alternatively, the missed last session method makes no assumption about the rationale underlying a client's unilateral decision to withdraw from services. Under this criterion, clients who drop out after the second session are indistinguishable from clients who fail to return after session 15 (Swift & Greenberg, 2015). The accuracy and utility of this assumption is questionable as higher "doses" (e.g., total number of sessions attended) are expected to produce larger therapeutic effects (Pfund et al., 2018). Given such concerns, the present investigation used the CSC method to classify PT. Under this procedure, the determination of dropout depends on whether clients unilaterally discontinue therapy before experiencing a meaningful reduction in the presenting problem, at a magnitude which cannot be attributed to chance. By anchoring the definition of dropout onto improvement, the CSC method produces estimates which are more likely to converge upon the intended construct (Swift & Greenberg, 2015), while simultaneously diverging from those previously reported in the literature. Thus, the contributions of this study have direct implications for the design and implementation of future trials. Implying it may be unnecessary for investigators to build empirically supported retention techniques, such as role induction or duration and patterns of change education, into the treatment protocol for future studies (Swift et al., 2012).

Further, it is difficult to place the observed dropout rate in the current study within the context of the empirical literature. First, the rate of PT is known to vary widely across treatment settings with estimates ranging anywhere from 0 to 74.2% (Swift & Greenberg, 2012). Also, data specific to PT in counselor training clinics are rarely published in professional journals and, when such data are available, researchers tend to operationalize PT using different classification criteria including therapist-determination or the missed last session method (Harris et al., 2020; Lampropoulos et al., 2009). To ensure comparability across studies, it is critical for researchers to thoroughly describe how PT was measured.

In the present sample, results suggest that both groups began VCP with comparable levels of symptom severity and appeared to demonstrate a trajectory of improvement. These findings suggest that VCP was unlikely to result in the reliable worsening of severity ratings. These findings are interpreted to provide support for the safety of VCP and to provide justifications for continued research into the effectiveness of this delivery method when operated by trainee clinicians. Results from the longitudinal growth model also appeared to suggest slightly faster improvement for clients in the VCP group. Perhaps one explanation is that clients in the VCP group experienced a consistent modality of intervention. Another possibility could be due to the methods used during the analysis.

The estimated deterioration rate was approximately 4% in the current sample. This estimate is consistent, albeit slightly lower than other studies of IP delivery. For example, in a recent meta-analysis of 228 clinical trials, the authors observed a deterioration rate of approximately 5% in the treatment group, compared with a rate of 7–13% across control conditions (Cuijpers et al., 2021). Similarly, large observational studies in community settings have also reported an overall deterioration rate of approximately 5% (Saxon et al., 2017; Rousmaniere et al., 2016). In training clinics, previous studies have reported estimates ranging from around 5.7% (Edmondstone et al., 2022) to nearly 10% (Carr et al., 2017). Although such findings lend support to the safety of VCP in this context, more research is needed to determine whether deterioration is the result of either side effects or malpractice and the extent to which negative effects occur in other domains beyond symptom severity (Herzog et al., 2019).

## Limitations

Results from the model fitting procedure should be interpreted with caution as the observed parameter estimates may be unstable due, in part, to the relatively small sample. Moreover, the present study did not include a specific measure of the therapeutic alliance. As such, future research should directly explore how the congruence/incongruence of client and counselor ratings of the alliance influence clinical outcomes when VCP is delivered by trainees.

One strength of this study is that both groups participated in VCP. This design feature makes it possible to explore how clients in the IP group reacted to switching from in-person to remote services in terms of both credibility and expectations. However, this design feature also represents a limitation as neither group was matched by start date or duration. Although growth models can accommodate for certain irregularities (e.g., flexible start times, missing observations, etc.), data from the IP group were characterized by sparse observations toward the beginning of counseling and dense observations toward the end. Data from the VCP group were the opposite. Therefore, it remains unknown whether differences

in the rate of change between groups are substantive or a statistical artifact.

## CONCLUSION


Results from the present investigation provide preliminary evidence for the safety of VCP when delivered in the training clinic environment. While a substantial portion of clients in the current study declined to switch from IP to VCP services, on average those who did had similar outcomes to clients in the IP group. In fact, the majority of clients in both groups reported lessening or stable depressive symptoms, and only a small percentage reported deterioration, at a rate slightly lower than what has been identified in previous IP samples. Such findings support the likelihood that clients are unlikely to deteriorate in VCP at a rate higher than routinely expected during the normal delivery of counseling services. Moreover, an unnecessary prolongation of suffering is also unlikely in cases presenting with mild to moderate severity ratings. However, it still remains unknown how clients with more severe and complex symptom profiles may respond to VCP in this clinical context. Finally, based on the low rate of PT, it appears reasonable to infer that clients view VCP as an acceptable delivery format within the context of a UBCTC.

## CONFLICT OF INTEREST

We have no known conflicts of interest to disclose.

## ORCID

Sean B. Hall  <https://orcid.org/0000-0002-2307-8324>

Annemarie Connor  <https://orcid.org/0000-0001-6583-8873>

Suzanne M. Dugger  <https://orcid.org/0000-0003-2730-8403>

Krista Casazza  <https://orcid.org/0000-0002-2018-2476>

## REFERENCES

- AEgisdottir, S., White, M., Spengler, P., Maugherman, A., Anderson, L., Cook, R., Nichols, C., Lampropoulos, G., Walker, B., Cohen, G., & Rush, J. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *Counseling Psychologist, 34*(3), 341–382. <https://doi.org/10.1177/0011000005285875>
- Al-Jabari, R., Murrell, A., Callahan, J., Cox, R., & Lester, E. (2019). Do distress level and waitlists impact termination in a training clinic? *Training and Education in Professional Psychology, 13*(2), 127–137. <https://doi.org/10.1037/tep0000223>
- American Counseling Association. (2014). *ACA code of ethics*. Alexandria, VA: American Counseling Association. <https://www.counseling.org/resources/aca-code-of-ethics.pdf>
- Barlow, D. H. (2010). Negative effects from psychological treatments: A perspective. *American Psychologist, 65*(1), 13–20. <https://doi.org/10.1037/a0015643>
- Batastini, A. B., Paprzycki, P., Jones, A. C. T., & MacLean, N. (2021). Are videoconferenced mental and behavioral health services just as good as in-person? A meta-analysis of a fast-growing practice. *Clinical Psychology Review, 83*, 101944. <https://doi.org/10.1016/j.cpr.2020.101944>
- Berryhill, M. B., Culmer, N., Williams, N., Halli-Tierney, A., Betancourt, A., Roberts, H., & King, M. (2019). Videoconferencing psychotherapy and depression: A systematic review. *Telemedicine and E-Health, 25*(6), 435–446. <https://doi.org/10.1089/tmj.2018.0058>
- Berryhill, M. B., Halli-Tierney, A., Culmer, N., Williams, N., Betancourt, A., King, M., & Ruggles, H. (2019). Videoconferencing psychological therapy and anxiety: A systematic review. *Family Practice, 36*(1), 53–63. <https://doi.org/10.1093/fampra/cmz072>
- Bowes, S., Ammirati, R., Costello, T., Basterfield, C., & Lilienfeld, S. (2020). Cognitive biases, heuristics, and logical fallacies in clinical practice: A brief field guide for practicing clinicians and supervisors. *Professional Psychology: Research and Practice, 51*(5), 435–445. <https://doi.org/10.1037/pro0000309>
- Brewin, C. R., & Bradley, C. (1989). Patient preferences and randomised clinical trials. *BMJ (Clinical Research Ed.), 299*(6694), 313–315. <https://doi.org/10.1136/bmj.299.6694.313>
- Callahan, J., Gustafson, S., Misner, J., Paprocki, C., Sauer, E., Saules, K., Schwartz, J., Swift, J., Whiteside, D., Wierda, K., & Wise, E. (2014). Introducing the association of psychology training clinics' collaborative research network: A study on client expectancies. *Training and Education in Professional Psychology, 8*(2), 95–104. <https://doi.org/10.1037/tep0000047>
- Callahan, J., & Watkins, C. (2018). The science of training II: Prepracticum and practicum training. *Training and Education in Professional Psychology, 12*(4), 231–244. <https://doi.org/10.1037/tep0000209>
- Carr, M. M., Saules, K. K., Koch, E. I., & Waltz, T. J. (2017). Testing the dose-response curve in a training clinic setting: Use of client pretreatment factors to minimize bias in estimates. *Training and Education in Professional Psychology, 11*(1), 26–32. <http://doi.org/10.1037/tep0000135>
- Choi, S., Schalet, B., Cook, K., & Cella, D. (2014). Establishing a common metric for depressive symptoms: Linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. *Psychological Assessment, 26*(2), 513–527. <https://doi.org/10.1037/a0035768>
- Constantino, M., Coyne, A., Boswell, J., Iles, B., & Visla, A. (2018). A meta-analysis of the association between patients' early perception of treatment credibility and their posttreatment outcomes. *Psychotherapy, 55*(4), 486–495. <https://doi.org/10.1037/pst0000168>
- Constantino, M., Visla, A., Coyne, A., & Boswell, J. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. *Psychotherapy, 55*(4), 473–485. <https://doi.org/10.1037/pst0000169>
- Cuijpers, P., Karyotaki, E., Ciharova, M., Miguel, C., Noma, H., & Furukawa, T. (2021). The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: A meta-analysis. *Acta Psychiatrica Scandinavica, 144*(3), 288–299. <https://doi.org/10.1111/acps.13335>
- Delgado, J., Kellett, S., Ali, S., McMillan, D., Barkham, M., Saxon, D., Donohoe, G., Stonebank, H., Mullaney, S., Eschoe, P., Thwaites, R., & Lucock, M. (2016). A multi-service practice research network study of large group psychoeducational cognitive behavioural therapy. *Behaviour Research and Therapy, 87*, 155–161. <https://doi.org/10.1016/j.brat.2016.09.010>
- Devilly, G., & Borkovec, T. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 31*(2), 73–86. [https://doi.org/10.1016/S0005-7916\(00\)00012-4](https://doi.org/10.1016/S0005-7916(00)00012-4)
- Dimidjian, S., & Hollon, S. (2010). How would we know if psychotherapy were harmful? *American Psychologist, 65*(1), 21–33. <https://doi.org/10.1037/a0017299>
- Dyason, K. M., Shanley, D. C., Hawkins, E., Morrissey, S. A., & Lambert, M. J. (2019). A systematic review of research in psychology training clinics: How far have we come? *Training and Education in Professional Psychology, 13*(1), 4–20. <https://doi.org/10.1037/tep0000196>
- Edmondstone, C., Pascual-Leone, A., Soucie, K., & Kramer, U. (2022). Therapist effects on outcome: Meaningful differences exist early in training. *Training and Education in Professional Psychology*, No Pagination Specified-No Pagination Specified. <https://doi.org/10.1037/tep0000402>

- Erekson, D., Janis, R., Bailey, R., Cattani, K., & Pedersen, T. (2017). A longitudinal investigation of the impact of psychotherapist training: Does training improve client outcomes? *Journal of Counseling Psychology, 64*(5), 514–524. <https://doi.org/10.1037/cou0000252>
- Eldridge, S. M., Lancaster, G. A., Campbell, M. J., Thabane, L., Hopewell, S., Coleman, C. L., & Bond, C. M. (2016). Defining feasibility and pilot studies in preparation for randomised controlled trials: Development of a conceptual framework. *Plos One, 11*(3), e0150205. <https://doi.org/10.1371/journal.pone.0150205>
- Essau, C., de la Torre-Luque, A., Lewinsohn, P., & Rohde, P. (2020). Patterns, predictors, and outcome of the trajectories of depressive symptoms from adolescence to adulthood. *Depression and Anxiety, 37*(6), 565–575. <https://doi.org/10.1002/da.23034>
- Finsrud, I., Nissen-Lie, H. A., Vrabell, K., Høstmælingen, A., Wampold, B. E., & Ulvenes, P. G. (2021). It's the therapist and the treatment: The structure of common therapeutic relationship factors. *Psychotherapy Research, 1*–12. <https://doi.org/10.1080/10503307.2021.1916640>
- Firth, N., Saxon, D., Stiles, W., & Barkham, M. (2020). Therapist effects vary significantly across psychological treatment care sectors. *Clinical Psychology & Psychotherapy, 27*(5), 770–778. <https://doi.org/10.1002/cpp.2461>
- Gallo, K. P., Comer, J. S., & Barlow, D. H. (2013). Single-case experimental designs and small pilot trial designs. In *The Oxford handbook of research strategies for clinical psychology*. (pp. 24–39). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199793549.001.0001>
- Gerstenblith, J. A., Kline, K. V., Hill, C. E., & Kivlighan Jr., D. M. (2021). The triadic effect: Associations among the supervisory working alliance, therapeutic working alliance, and therapy session evaluation. *Journal of Counseling Psychology*, Advanced online publication. <https://doi.org/10.1037/cou0000567>
- Gruber, J., Prinstein, M. J., Clark, L. A., Rottenberg, J., Abramowitz, J. S., Albano, A. M., Aldao, A., Borelli, J. L., Chung, T., Davila, J., Forbes, E. E., Gee, D. G., Hall, G. C. N., Hallion, L. S., Hinshaw, S. P., Hofmann, S. G., Hollon, S. D., Joormann, J., Kazdin, A. E., ..., & Weinstock, L. M. (2021). Mental health and clinical psychological science in the time of COVID-19: Challenges, opportunities, and a call to action. *American Psychologist, 76*(3), 409–426. <https://doi.org/10.1037/amp0000707>
- Guo, B., Kaylor-Hughes, C., Garland, A., Nixon, N., Sweeney, T., Simpson, S., Dalgleish, T., Ramana, R., Yang, M., & Morriss, R. (2017). Factor structure and longitudinal measurement invariance of PHQ-9 for specialist mental health care patients with persistent major depressive disorder: Exploratory structural equation modelling. *Journal of Affective Disorders, 219*, 1–8. <https://doi.org/10.1016/j.jad.2017.05.020>
- Hames, J. L., Bell, D. J., Perez-Lima, L. M., Holm-Denoma, J. M., Rooney, T., Charles, N. E., Thompson, S. M., Mehlenbeck, R. S., Tawfik, S. H., Fondacaro, K. M., Simmons, K. T., & Hoersting, R. C. (2020). Navigating uncharted waters: Considerations for training clinics in the rapid transition to telepsychology and telesupervision during COVID-19. *Journal of Psychotherapy Integration, 30*(2), 348–365. <https://doi.org/10.1037/int0000224>
- Hardy, G. E., Barkham, M., Shapiro, D. A., Reynolds, S., Rees, A., & Stiles, W. B. (1995). Credibility and outcome of cognitive—Behavioural and psychodynamic—Interpersonal psychotherapy. *British Journal of Clinical Psychology, 34*(4), 555–569. <https://doi.org/10.1111/j.2044-8260.1995.tb01489.x>
- Harris, S., Mitchell, M., Tabet, S., & Hundley, G. (2020). Predictors of client attrition in a university-based community counseling clinic. *Journal of Counseling and Development, 98*(1), 74–82. <https://doi.org/10.1002/jcad.12301>
- Hatchett, G., & Park, H. (2003). Comparison of four operational definitions of premature termination. *Psychotherapy, 40*(3), 226–231. <https://doi.org/10.1037/0033-3204.40.3.226>
- Hatfield, D., McCullough, L., Frantz, S., & Krieger, K. (2010). Do we know when our clients get worse? an investigation of therapists' ability to detect negative client change. *Clinical British Journal of Psychiatry Psychology & Psychotherapy, 17*(1), 25–32. <https://doi.org/10.1002/cpp.656>
- Herzog, P., Lauff, S., Rief, W., & Brakemeier, E. L. (2019). Assessing the unwanted: A systematic review of instruments used to assess negative effects of psychotherapy. *Brain and behavior, 9*(12), e01447. <https://doi.org/10.1002/brb3.1447>
- Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75*(4), 800–802. <https://doi.org/10.1093/biomet/75.4.800>
- Howard, L., & Thornicroft, G. (2006). Patient preference randomised controlled trials in mental health research. *The British Journal of Psychiatry, 188*, 303–304. <https://doi.org/10.1192/bjp.188.4.303>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. <https://doi.org/10.1037/0022-006X.59.1.12>
- Kroenke, K., Spitzer, R., & Williams, J. (2001). The PHQ-9 – Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Lampropoulos, G., Schneider, M., & Spengler, P. (2009). Predictors of early termination in a university counseling training clinic. *Journal of Counseling and Development, 87*(1), 36–46. <https://doi.org/10.1002/j.1556-6678.2009.tb00547.x>
- Levis, B., Benedetti, A., Thombs, B., Akena, D., Arroll, B., Ayalon, L., Azar, M., Baradaran, H., Baron, M., Bombardier, C., Boruff, J., Butterworth, P., Carter, G., Chagas, M., Chan, J., Chiovitti, M., Clover, K., Conwell, Y., & Cuijpers, P., ... depression Screening Data DEPRESSD. (2019). Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: Individual participant data meta-analysis. *BMJ-British Medical Journal, 365*, 11476. <https://doi.org/10.1136/bmj.11476>
- Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2*(1), 53–70. <https://doi.org/10.1111/j.1745-6916.2007.00029.x>
- Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods, 52*(2), 464–488. <https://doi.org/10.3758/s13428-019-01246-w>
- Mallah, S. I., Ghorab, O. K., Al-Salmi, S., Abdellatif, O. S., Tharmaratnam, T., Iskandar, M. A., Sefen, J. A. N., Sidhu, P., Atallah, B., El-Lababidi, R., & Al-Qahtani, M. (2021). COVID-19: Breaking down a global health crisis. *Annals of Clinical Microbiology and Antimicrobials, 20*(1), 35. <https://doi.org/10.1186/s12941-021-00438-7>
- Manthei, R. J., Vitalo, R. L., & Ivey, A. E. (1982). The effect of client choice of therapist on therapy outcome. *Community Mental Health Journal, 18*(3), 220–229. <https://doi.org/10.1007/BF00754338>
- Matsumoto, K., Hamatani, S., & Shimizu, E. (2021). Effectiveness of videoconference-delivered cognitive behavioral therapy for adults with psychiatric disorders: Systematic and meta-analytic review. *Journal of Medical Internet Research, 23*(12), e31293. <https://doi.org/10.2196/31293>
- Memillan, D., Gilbody, S., & Richards, D. (2010). Defining successful treatment outcome in depression using the PHQ-9: A comparison of methods. *Journal of Affective Disorders, 127*(1–3), 122–129. <https://doi.org/10.1016/j.jad.2010.04.030>
- Musliner, K., Munk-Olsen, T., Eaton, W., & Zandi, P. (2016). Heterogeneity in long-term trajectories of depressive symptoms: Patterns, predictors and outcomes. *Journal of Affective Disorders, 192*, 199–211. <https://doi.org/10.1016/j.jad.2015.12.030>
- Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy, 55*(4), 303–315. <https://doi.org/10.1037/pst0000193>
- Norman, G. R., Grierson, L. E. M., Sherbino, J., Hamstra, S. J., Schmidt, H. G., & Mamede, S. (2018). Expertise in medicine and surgery. In A. M. Williams, A. Kozbelt, K. A. Ericsson, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance*. (2nd edn., pp. 331–355). Cambridge University Press; Cambridge Core. <https://doi.org/10.1017/9781316480748.019>
- Norwood, C., Moghaddam, N. G., Malins, S., & Sabin-Farrell, R. (2018). Working alliance and outcome effectiveness in videoconferencing psychotherapy: A systematic review and noninferiority meta-analysis.

- Clinical Psychology & Psychotherapy*, 25(6), 797–808. <https://doi.org/10.1002/cpp.2315>
- Patel, J., Oh, Y., Rand, K., Wu, W., Cyders, M., Kroenke, K., & Stewart, J. (2019). Measurement invariance of the patient health questionnaire-9 (PHQ-9) depression screener in US adults across sex, race/ethnicity, and education level: NHANES 2005–2016. *Depression and Anxiety*, 36(9), 813–823. <https://doi.org/10.1002/da.22940>
- Pfund, R., Peter, S., Whelan, J., & Meyers, A. (2018). When does premature treatment termination occur? Examining session-by-session dropout among clients with gambling disorder. *Journal of Gambling Studies*, 34(2), 617–630. <https://doi.org/10.1007/s10899-017-9733-z>
- Richardson, E., & Richards, J. (2008). Factor structure of the PHQ-9 screen for depression across time since injury among persons with spinal cord injury. *Rehabilitation Psychology*, 53(2), 243–249. <https://doi.org/10.1037/0090-5550.53.2.243>
- Rousmaniere, T., Swift, J., Babins-Wagner, R., Whipple, J., & Berzins, S. (2016). Supervisor variance in psychotherapy outcome in routine practice. *Psychotherapy Research*, 26(2), 196–205. <https://doi.org/10.1080/10503307.2014.963730>
- Runciman, W., Hibbert, P., Thomson, R., Van Der Schaaf, T., Sherman, H., & Lewalle, P. (2009). Towards an international classification for patient safety: Key concepts and terms. *International Journal for Quality in Health Care*, 21(1), 18–26. <https://doi.org/10.1093/intqhc/mzn057>
- Saxon, D., Barkham, M., Foster, A., & Parry, G. (2017). The contribution of therapist effects to patient dropout and deterioration in the psychological therapies. *Clinical Psychology & Psychotherapy*, 24(3), 575–588. <https://doi.org/10.1002/cpp.2028>
- Schaeffer, N., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology*, 29, 65–88. <https://doi.org/10.1146/annurev.soc.29.110702.110112>
- Sekhon, M., Cartwright, M., & Francis, J. J. (2017). Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework. *BMC Health Services Research*, 17(1), 88. <https://doi.org/10.1186/s12913-017-2031-8>
- Spengler, P., & Pilipis, L. (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>
- Steinberg, L., & Rogers, A. (2020). Changing the scale: The effect of modifying response scale labels on the measurement of personality and affect. *Multivariate Behavioral Research*, <https://doi.org/10.1080/00273171.2020.1807305>
- Swift, J., & Callahan, J. (2008). A delay discounting measure of great expectations and the effectiveness of psychotherapy. *Professional Psychology: Research and Practice*, 39(6), 581–588. <https://doi.org/10.1037/0735-7028.39.6.581>
- Swift, J., Callahan, J., Cooper, M., & Parkin, S. (2018). The impact of accommodating client preference in psychotherapy: A meta-analysis. *Journal of Clinical Psychology*, 74(11), 1924–1937. <https://doi.org/10.1002/jclp.22680>
- Swift, J., Callahan, J., & Levine, J. (2009). Using clinically significant change to identify premature termination. *Psychotherapy*, 46(3), 328–335. <https://doi.org/10.1037/a0017003>
- Swift, J., & Greenberg, R. (2012). Premature Discontinuation in Adult Psychotherapy: A Meta-Analysis. *Journal of Consulting and Clinical Psychology*, 80(4), 547–559. <https://doi.org/10.1037/a0028226>
- Swift, J. K., & Greenberg, R. P. (2015). *Premature termination in psychotherapy: Strategies for engaging clients and improving outcomes*. American Psychological Association. <https://doi.org/10.1037/14469-000>
- Swift, J., Greenberg, R., Tompkins, K., & Parkin, S. (2017). Treatment refusal and premature termination in psychotherapy, pharmacotherapy, and their combination: A meta-analysis of head-to-head comparisons. *Psychotherapy*, 54(1), 47–57. <https://doi.org/10.1037/pst0000104>
- Swift, J. K., Greenberg, R. P., Whipple, J. L., & Kominiak, N. (2012). Practice recommendations for reducing premature termination in therapy. *Professional Psychology: Research and Practice*, 43(4), 379–387. <https://doi.org/10.1037/a0028291>
- Wampold, B. (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). *Great psychotherapy debate: The evidence for what makes psychotherapy work*. 2nd edn. Routledge <https://doi.org/10.1007/978-3-319-02603-9>
- Wellck, S. (2010). *Testing statistical hypotheses of equivalence and non-inferiority*. 2nd edn. Chapman and Hall/CRC. <https://doi.org/10.1201/EBK1439808184>
- Wilcox, R., Erceg-Hurn, D., Clark, F., & Carlson, M. (2014). Comparing two independent groups via the lower and upper quantiles. *Journal of Statistical Computation and Simulation*, 84(7), 1543–1551. <https://doi.org/10.1080/00949655.2012.754026>
- Windle, E., Tee, H., Sabitova, A., Jovanovic, N., Priebe, S., & Carr, C. (2020). Association of patient treatment preference with dropout and clinical outcomes in adult psychosocial mental health interventions a systematic review and meta-analysis. *JAMA Psychiatry*, 77(3), 294–302. <https://doi.org/10.1001/jamapsychiatry.2019.3750>

**How to cite this article:** Hall, S. B., Bartley, A. G., Wenk, J., Connor, A., Dugger, S. M., & Casazza, K. (2022). Rapid transition from In-person to videoconferencing psychotherapy in a counselor training clinic: A safety and feasibility study during the COVID-19 pandemic. *Journal of Counseling & Development*, 1–12. <https://doi.org/10.1002/jcad.12439>