UNIVERSITY OF CALIFORNIA

Los Angeles

Application of Organizational Management Concepts to the Study of Clinical Supervision in Community Mental Health: The Influence of Clinical Supervision in the Therapy Room and Broader Organization

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy in Psychology

by

Meredith Rose Boyd

2023

© Copyright by

Meredith Rose Boyd

2023

ABSTRACT OF THE DISSERTATION

Application of Organizational Management Concepts to the Study of Clinical Supervision in Community Mental Health: The Influence of Clinical Supervision in the Therapy Room and Broader Organization

by

Meredith Rose Boyd

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2023

Professor Bruce Fredrick Chorpita, Chair

Although there is a substantial evidence base of interventions to treat child and adolescent mental health conditions, these interventions remain underused in community mental health settings. Increasingly, clinical supervision has gained attention as an opportunity for promoting use of interventions supported by research. At the interpersonal level, supervisors are well-positioned to support ongoing skill development of therapists through such strategies as role plays, skill modeling, and performance feedback. Supervisors are also well-positioned to impact therapists' experience of the organizations in which they work. This is particularly important given the high stress, burnout, and organizational turnover characteristic of community mental health agencies. Specifically, supervisors are "middle managers," interfacing both with upper-level management and frontline workers. This position affords the opportunities to facilitate communication from the top down and bottom up.

Acknowledging the potential influence of supervisors on the organizations in which they work and on the individual therapists they oversee, this dissertation examined supervisor impact on therapist experience of their work environment (i.e., organizational climate) and on therapist use of practices supported by research. The first study aimed to examine the role of clinical supervisors as drivers of therapist perceptions of organizational climate within supervisory teams. An understanding of the potential impact of the supervisor on therapist perceptions of organizational climate could inform additional interventions to positively influence climate and facilitate improvement in the quality of care delivered. The second study examined supervision strategies that create a sense of accountability informed by a model from the organizational management literature in which accountability is fostered by behaviors that strengthen links between *prescriptions* (directives to perform a behavior), *events* (contexts in which the behavior can be performed) and *identity* (personal qualities and skills). Informed by this model, the second study aimed to identify the presence of supervision accountability strategies and investigate the association between accountability strategies and therapist follow through use of a therapeutic practice identified during supervision in the subsequent treatment session. As a first step in capitalizing on supervision to improve client outcomes, it is important to determine *how* supervision influences therapist behavior during treatment by directly testing supervision strategies.

The dissertation of Meredith Rose Boyd is approved.

Kimberly D. Becker

Alayna Park

Denise A. Chavira

Lauren Christina Ng

Bruce Fredrick Chorpita, Committee Chair

University of California, Los Angeles

2023

Table of Contents

Chapter 1 ..... *1*

Abstract 2

Introduction 4
- Organizational Climate 5
- Climate Strength 5
- Levels of Analysis 6
- Middle Managers 6
- Study Aims 8

Methods 10
- Participants 10
- Measures 11
- Data Analysis 16

Results 19
- Aim 1 19
- Aim 2 20
- Aim 3 21

Discussion 22
- Aim 1: Perceptions of Organizational Climate 22
- Aim 2: Predictors of Climate 26
- Aim 3: Climate as a Predictor of Youth Engagement 27
- Limitations and Future Directions 28
- Conclusions 30

Tables 31

References 41

Chapter 2 ..... *50*

Abstract 51

Introduction 53
- Accountability in Supervision 53
- Prescription-Event Link 56
- Identity-Prescription Link 57
- Identity-Event Link 59
- Study Aims 60

Methods 61
- Participants 61
- Behavioral Observational Coding Systems 62
- ACE-BOCS 63
- URECA-E 65
- Data Analysis 66

Results 68

Aim 1 68
Aim 2 68
Discussion 70
Prescription-Event Link 71
Identity-Prescription Link 72
Event-identity link 72
Limitations 73
Future Directions 75
Conclusions 77
Summative Discussion of Dissertation Findings *77*
Figures 80
Tables 82
References 86

Figures and Table

Table 1. Demographics and Supervisors and Therapists (Study 1)

Table 2. Supervision Group Descriptive Statistics, Intraclass Correlation Coefficients, and Reliability Estimates for Climate Subscales and Total Score (Study 1)

Table 3. Agreement Index (a_{wg}) by Supervision Group (Study 1)

Table 4. Predictors of Therapist Perception of Autonomy in Organization (Study 1)

Table 5. Predictors of Therapist Perception of Autonomy in Organization (Study 1)

Table 6. Predictors of Therapist Perception of Organizational Climate (Study 1)

Table 7. Supervision Team Autonomy and Climate Strength Predicting Youth Baseline Treatment Engagement (Study 1)

Table 8. Supervision Team Cohesion and Climate Strength Predicting Youth Baseline Treatment Engagement (Study 1)

Table 9. Supervision Team Climate Total Score and Climate Strength Predicting Youth Baseline Treatment Engagement (Study 1)

Table 10. Demographics of Supervisor and Therapists (Study 2)

Table 11. Agreement and Reliability Estimates for Supervision and Treatment Actions (Study 2)

Table 12. Effect of Presence of Supervision Behaviors on Therapist use of Expected Practice (Study 2)

Table 13. Effect of Count of Supervisor behaviors on Therapist use of Expected Practice (Study 2)

Figure 1. ACE-BOCS Codes Mapped onto Accountability Links (Study 2)

Figure 2. ACE-BOCS to URECA-E Mapping (Study 2)

Vita

EDUCATION

2023	Ann Arbor VA Healthcare System Pre-doctoral Internship
2016	University of California, Los Angeles Master of Arts in Psychology
2015	Indiana University, Bloomington Bachelor of Arts in Psychology, Departmental Honors

SELECTED AWARDS

2023	Michael J. Goldstein Distinguished Dissertation Award
2018, 2020	UCLA Graduate Summer Research Mentorship Fellowship
2015	James A. Dinsmoor Outstanding Psychology Research Award Recipient

SELECTED PUBLICATIONS

Boyd, M. R., Lewis. C. C., Marti, N., Albright, K. (2023) Mediators of measurement-based care implementation in community mental health settings: Results from a mixed methods evaluation. *Implementation Science. 17*(1), 1-18. https://doi.org/10.1186/s13012-022-01244-1

Lewis, C. C., Marti, C. N., Scott, K. S., Walker, M., Boyd, M. R., Puspitasari, A., Mendel, P., & Kroenke, K., (2022). Standardized versus tailored implementation of measurement based care for depression in community mental health centers: A cluster randomized trial. *Psychiatric Services.* https://doi.org/10.1176/appi.ps.202100284

Hanano, M., Rith-Najarian, L., Boyd, M. R., & Chavira, D. (2022). Measuring Adherence Within a Self-Guided Online Intervention for Depression and Anxiety: Secondary Analyses of a Randomized Controlled Trial. *JMIR Mental Health*, *9*(3), e30754. https://doi.org/10.2196/30754

Albright, K., Navarro, E. I., Jarad, I., Boyd, M. R., Powell, B., & Lewis, C. C. (2021). Communication strategies to facilitate the implementation of new clinical practices: A qualitative study of community mental health therapists. *Translational Behavioral Medicine. 12*(2), 324-334 https://doi.org/10.1093/tbm/ibab139

Boyd, M. R., Park, A. L., Becker, K. D., & Chorpita, B. F. (2021). The relation between training asymmetry and supervisory working alliance: Implications for the role of supervisors in implementation. *The Clinical Supervisor*, *40*(1) 1-19. https://doi.org/10.1080/07325223.2020.1871460

Lewis, C. C., Boyd, M. R., Walsh-Bailey, C., Lyon, A. R., Beidas, R., Mittman, B., Aarons, G. A., & Weiner, B. J. (2020). A systematic review of empirical studies examining mechanisms of implementation in health. *Implementation Science, 15*(1), 1-25. https://doi.org/10.1186/s13012-020-00983-3

Keenan-Miller, D., Boyd, M. R., Westman, J., & Chorpita, B. F. (2020). Setting standards for supervision in a clinical science training clinic. The *Behavior Therapist.* 43(4), 127-133.

Lewis, C. C., Boyd, M. R., Puspitasari, A., Navarro, E., Howard, J., Kassab, H. … Kroenke K. (2019). An implementation-science informed narrative review of measurement-based care in behavioral health. *JAMA Psychiatry, 76*(3), 324-335. https://doi.org/10.1001/jamapsychiatry.2018.3329

Boyd, M. R., Powell, B. J., Endicott, D. & Lewis, C. C. (2017). A method for tracking implementation strategies: An exemplar implementing measurement-based care in community behavioral health clinics. *Behavior Therapy, 49*(4), 525-537. https://doi.org/10.1016/j.beth.2017.11.012

Boyd, M. R., Lewis, C. C., Scott, K., Krendl, A., & Lyon, A. R. (2017). The creation and validation of a measure of effective attributes of trainers (MEAT). *Implementation Science*, *12*(1), 73. https://doi.org/10.1016/j.beth.2017.11.012

Chapter 1

The Impact of Supervision Team on Therapist Perception of Organizational Climate

Abstract

In the context of mental health service organizations, positive organizational climate — employee perceptions of their work environment and the impact of this environment on well-being and functioning — has been found to be associated with a number of desirable organizational and client-level outcomes. Clinical supervisors are well-positioned in community mental health organizations to impact climate positively, as they serve as intermediaries between higher-level administrators who drive the policies and procedures and the therapists impacted by such decisions. The present study examined the role of clinical supervisors as drivers of therapist perceptions of organizational climate within supervisory teams. Specifically, the present study investigated: (1) shared perceptions of organizational climate among therapists on the same supervisory team; (2) predictors of therapist climate perceptions; (3) association between supervisor-level organizational climate and client treatment engagement.

Data were collected from 86 school-based therapists overseen by 22 supervisors. To accomplish the first aim, indices of interrater agreement of therapists on the same supervisory team ($r_{wg(j)}$ and a_{wg}) and indices of interrater reliability (ICC (1) and ICC (2)) were examined to test the hypothesis that therapists on the same supervisory team shared more similar perceptions of climate compared with therapists on different supervisory teams. Regarding the second aim, multi-level models were used to examine the associations between supervisor attitudes towards evidenced-based practices and therapist perceptions of supervisor communication and organizational climate. To accomplish the third aim, multi-level models were used to examine the association between supervisor team-level aggregation of organizational climate (i.e., mean score) and client self-report of treatment engagement moderated by climate strength (i.e., variability in perceptions of climate from members of the same team as measured by standard

deviation). Results yielded evidence of perceptions of organizational cohesion and autonomy that were shared among therapists from the same supervision team and distinct from therapists on different supervision teams. Therapist perception of their supervisor's communication was positively associated with perceptions of organizational cohesion and autonomy. These results build upon past research describing the key role of middle managers as intermediaries between upper-level management and front-line workers. Further elucidating the role of supervisors in mental health organizations will be important for maximizing workforce capacity to address the substantial and growing demand for quality mental health treatment.

Introduction

There is a substantial evidence base of interventions to treat child and adolescent mental health conditions (Chorpita et al., 2011). However, interventions supported by research remained underused in community mental health settings where the majority of children and families access services (Weisz et. al., 2006; Borntrager et al., 2013; Garland et al., 2010). Differences in the organizational context in which evidence-based treatments (EBTs) are developed versus that in which their use is intended has been hypothesized as contributing to the gap between what is established in the evidence base and what is done in the therapy room (Hoagwood et al., 2001; Chorpita et al., 2014; Southam-Gerow et al., 2008). For example, therapists in community mental health settings experience significant strain and competing demands within their organizations distinct from therapists who participate in EBT efficacy trials. Specifically, study therapists are typically associated with high-resource research institutes, treat a small number of clients, and receive a high degree of support and supervision (Hoagwood et al., 2001). In contrast, community-based therapists often carry large clinical caseloads (e.g., Lau et al., 2018; Dorsey et al., 2017) and treat high acuity clients with greater comorbidity and adverse life circumstances compared with clients included in study trials (Chorpita et al., 2014; Southam-Gerow et al., 2008). Productivity standards, i.e., organizational mandates related to the frequency or duration of client contact, are also common in community mental health contexts and must be balanced with time-consuming documentation requirements (Franco, 2016). With these conditions in mind, it is perhaps unsurprising that stress, emotional exhaustion, and burnout among community-based therapists are high (Morse et al., 2012). These conditions are predictive of turnover and, indeed, turnover in community mental health organizations is estimated to be as high as 50% in a given year (Woltmann et. al., 2008).

Organizational Climate

Increasingly, organizational climate has gained attention as a determinant that may hinder or facilitate the implementation of EBTs in community mental health settings. Although there are several definitions in the literature, climate is commonly defined as employee perceptions of their work environment and the impact of this environment on well-being and functioning (Glisson & James, 2002; Ouellette et al., 2020). Positive organizational climate has been found to be associated with a number of desirable outcomes including: improved therapist attitudes towards their work and EBTs (e.g., Aarons & Sommerfeld, 2012; Brimhall et al., 2015; Aarons & Sawitzky, 2006), improved EBT quality and fidelity (e.g., Williams et al., 2019; Olin et al., 2014), increased client engagement and satisfaction with services (e.g., Greener et al., 2006; Lehman et al., 2002), and better clinical outcomes (e.g., Glisson & Green, 2011). Conversely, poor organizational climate is associated with therapist burnout and organizational turnover (Aarons & Sawitzky, 2006; Glisson et al., 2008). For example, Glisson et al. (2008) created profiles of organizational climates for 100 clinics across the United States and found that yearly turnover rates of organizations with the most negative climates were double the rates of organizations with the most positive climates. Such findings emphasize the importance of seeking to improve climate as a means for increasing organizational capacity of community mental health organizations to provide quality care.

Climate Strength

In addition to how positively or negatively employees view organizational climate, there is also some emerging evidence for the importance of how similarly employees view climate. Climate strength is defined as within-group variability in perceptions of climate and is commonly represented by the standard deviation of climate ratings at the organizational level

being examined (see "Levels of Analysis" section below; Schneider et al., 2012). Researchers such as Schneider et al. (2012) have hypothesized that the impact of climate on outcomes of interest is greater when perceptions of climate are more closely shared by all individuals in the organization. In the broader organizational management literature, there is some evidence to support this claim (e.g., Schneider et al., 2002; Colquitt et al., 2002; González-Romá et al., 2002) though this has yet to be examined in the context of community mental health agencies specifically.

Levels of Analysis

Availability, Responsiveness, and Continuity (ARC) is an example of an intervention that has been shown to improve climate and reduce staff turnover by intervening at the full organizational as well as the interorganizational level (i.e., collections of organizations and relevant stakeholders in the community; Glisson & Schoenwald, 2005; Glisson et al., 2006). Although ARC uses a macro-level approach, it is possible that intervening at more micro-levels of the organization might also yield changes in climate and be more pragmatic depending on the size and composition of the organization. Community mental health organizations often contain several organizational levels. For example, in school-based community mental health agencies, there is the level of the school district, regions or neighborhoods within a district, clinics within a region, care teams within clinics, and individuals within teams. This raises the following question: what level or levels of the organization should be targeted when seeking to improve organizational climate?

Middle Managers

A focus on middle managers (i.e., front-line leaders such as employees' direct supervisors) is one potentially feasible means for improving the climate of an organization

(Birken et al., 2012). Specifically, in community mental health settings, it is common for clinical supervisors to oversee the practice of therapists (Schoenwald et al., 2008). Clinical supervisors represent middle managers within mental health systems. They are well-positioned to positively impact organizational climate as they work directly with therapists who deliver treatment and serve as intermediaries between higher level administrators who dictate the policies and procedures and the therapists impacted by such decisions. Indeed, in their study of the activities performed by supervisors during supervision, Bailin et al. (2018) found that clinical supervisors spend time both on administrative tasks (i.e., implementation of policies and procedures) as well as supporting therapist clinical practice (e.g., case conceptualization, recommendation of practices, and modeling skills).

There is emerging evidence of the influence of supervisors in community mental health organizations (Aarons et al., 2011; Green et al., 2014; Brimhall et. al., 2016). For example, Aarons et al. (2011) found that provider's perceptions of the leadership style of their direct supervisor was associated with attitudes towards use of evidence-based practices (EBP). Further, in their study of clinical and case management providers and their direct supervisors in community mental health agencies in San Diego County, Green et al. (2014) found that strong leadership was associated with positive organizational climate which was, in turn, associated with working alliance between providers and their clients. However, in the reviewed studies, analyses were conducted by nesting providers within sites rather than within supervisor, making it difficult to determine the extent to which provider attitudes and perceptions were influenced by their supervisor rather than by other aspects of the organization. Because providers were asked to react to leadership specifically, it can be inferred that supervisors impacted provider perceptions. An alternative explanation is that, additional, unaccounted for organizational variables may have

influenced the way providers experienced the leadership of their supervisor. Finally, in their qualitative study of providers in a child welfare system experiencing changes in their positions due to implementation of a child assessment and response intervention, Bunger et al. (2019) found that supervisors shaped provider perceptions of organizational climate through their communication about the intervention, day-to-day support using the new intervention, and persuasion about the value of the intervention.

An understanding of the potential impact of the supervisor on therapist perceptions of organizational climate could inform additional interventions to improve climate, reduce organizational turnover, and enhance the quality of care delivered. If it is found that supervisors influence organizational climate, mental health organizations might prioritize training for supervisors to increase their skill and competence as team leaders or might protect additional time in the workday for supervisors to engage in tasks that support their supervisory team. Organizations might also monitor the organizational climate of supervisory teams and use this information to provide targeted support to supervisors of teams with poorer climates.

Study Aims

The overarching goal of the present study was to better understand the potential role of clinical supervisors as drivers of therapist perceptions of organizational climate and subsequent impact on client treatment engagement. The aims of this study were:

Aim 1: Determine if perceptions of organizational climate are shared among therapists in the same supervision teams and distinct from therapists on different teams.

- H1: Consistent with findings from climate studies from the organizational management literature, it is hypothesized that therapists in the same supervisory team will have more shared perceptions of organizational climate compared with therapists in different

supervisory teams as indicated by interrater agreement (a_{wg} and $r_{wg(j)}$) and interrater reliability (intraclass correlation coefficient, ICC; Schneider et al., 2013) calculated from multilevel models of therapists nested within supervisory teams.

Aim 2: Investigate candidate, supervisor predictors of therapist perception of climate.

- H2.1: Previous research suggests that therapist attitudes towards EBPs is influenced by organizational climate (Aarons & Sawitzky, 2006). Given these previously identified associations between climate and attitudes, the present study will investigate the association between supervisor attitudes and therapist climate perceptions. It is hypothesized that supervisor attitudes toward EBPs will be significantly and positively associated with therapist perceptions of organizational climate.
- H2.2: Informed by research about the role of middle managers in shaping communication from upper management to direct reports (Bailen et al., 2018; Bunger et al., 2019), it is hypothesized that therapist perceptions of supervisor communication will be significantly and positively associated with therapists' perception of organizational climate.

Aim 3: Determine if organizational climate scores aggregated at supervisory team level are predictive of client treatment engagement and moderated by climate strength (i.e., variability of perceptions as measured by standard deviation).

- H3: Informed by previous findings linking climate to client engagement (e.g., Greener et al., 2006; Lehman et al., 2002), it is hypothesized that teams with more positive organizational climate will have clients with greater treatment engagement. This relationship will be moderated by climate strength such that teams with less dispersion in climate perceptions (i.e., stronger climate) and more positive perceptions of climate will have clients with greater treatment engagement (Schneider et al., 2013).

Methods

Participants

Participants for the present study were recruited as part of the Reaching Families Study, a multi-site cluster randomized trial examining the impact of a coordinated knowledge system on therapist use of evidence and subsequent client engagement in school-based mental health services. Supervisors (n = 22) and therapists (n = 86) were included in the present study if they completed the main measure of interest (Texas Christian University Organizational Climate Scales, TCU-ORG; Lehman et al., 2001) and were part of pre-study supervision dyads (i.e., not a supervision dyad created for the purpose of study participation).

Supervisors and therapists were employed by either the Los Angeles Unified School District School Mental Health and Wellness Center Program in urban Los Angeles, CA (therapists n = 45, supervisors n = 12) or by the South Carolina Department of Mental Health in rural South Carolina (i.e., Pee Dee and Santee-Wateree catchment areas; Pee Dee therapists n = 19, supervisors n = 4; Santee-Wateree, therapists n = 22, supervisors n = 6). Supervisors and therapists were predominantly master's-level (supervisors 100%, therapists 98.11%) and women (supervisors 95.45%, therapists 90.70%). Supervisors were 50% Black/African American, 18.18% Hispanic/Latinx, 18.18% White/Caucasian. Therapists were 41.86% Black/African American, 39.53% Hispanic/Latinx and 15.20% White/Caucasian. See Table 1 for full demographic information for the study sample.

Excluded participants (n = 58) were compared to included participants (n = 108) across demographic variables to determine if there were statistically significant differences between these groups that might suggest bias in results. Welch's T-tests did not reveal differences in years of clinical experience and age. Chi-Square Goodness of Fit Tests did not reveal differences in the

proportion of White or Latinx providers in the excluded versus included participant samples. Across the full sample, the number of Asian providers was too small for Chi-Square testing (included $n = 4$; excluded $n = 4$). The proportion of Black providers in the excluded versus included sample was significantly different ($x^2(1, 166) = 5.53$, $p = .02$) such that the excluded participants had a smaller proportion of Black providers (22.06%) compared with the included participants (55.56%). In addition, the proportion of South Carolina providers in the excluded versus included sample was significantly different ($x^2(1, 166) = 16.44$, $p < .001$) such that the excluded participants had a smaller proportion of South Carolina providers (13.23%) compared with the included participants (47.22%).

Due to staff schedules and availability in California, some supervisory dyads were created for the purpose of the Reaching Families Study and did not exist prior. Supervisory dyad creation did not occur in South Carolina. As stated previously, study-created dyads were excluded from the present analyses resulting in a lower proportion of California providers in the included sample. Exclusion of study-created California supervisory dyads also resulted in a higher proportion of Black providers in the included sample as South Carolina had more Black providers compared with California. Across the full sample, the number of male providers was too small for Chi-Square testing (included $n = 9$; excluded $n = 2$).

Measures

Climate. Therapists completed the Texas Christian University Organizational Climate Scales (TCU-ORG; Lehman et al., 2001). The TCU-ORG is comprised of 30 self-report items that assess the following domains of organizational climate: clarity of organizational mission (e.g., “Your program operates with clear goals and objectives”; subscale labeled, “Mission”), cohesion among staff members (e.g., “Staff members at your program work together as a team”;

subscale labeled “Cohesion”), autonomy when making decisions (e.g., “Counselors in your program are given broad authority in treating their clients”; subscale labeled “Autonomy”), communication (e.g., “The formal and informal communication channels in your program work very well”; subscale labeled “Communication”), stress (e.g., “Staff members at your program often show signs of high stress and strain”; subscale labeled “Stress”), and attitudes toward change (e.g., “The general attitude in your program is to accept new and changing technology”; subscale labeled “Change”). Each item was scored on a four-point scale from 0 (strongly disagree) to 3 (strongly agree), with lower scores indicative of poorer climate. Consistent with the scoring instructions for this measure (Institute of Behavioral Research, 2005), scores for each subscale were calculated by taking the mean for the number of completed items and subscales were removed listwise if more than 50% of items were missing in the given subscale. In the current data set, 11.6% (n = 10) of participants had missing data and one subscale for one participant was missing more than 50% of the items for that subscale.

The wording of this measure of organizational climate aligns with Schneider’s (2013) definition of climate as *shared* perceptions about the *collective* impact of work environment given that respondents were asked to reflect on how they perceive the organization impacting other staff members (e.g., “*staff members* at your program often show signs of high stress and strain” rather than, “*I feel* a high degree of stress and strain). To reinforce the referent of the survey as the broader organization, the following instruction was included at the top of the survey in bold: “For the following 30 questions, please consider the School Mental Health Program.” Referent-shift consensus items refer to items that are worded such to prompt respondents to answer questions in relation to the same level of aggregation, in this specific case to answer questions in relation to the organization to which they belong rather than another level

of grouping such as the school system as a whole or the individual school in which they work (Chan 1998). In Lehman et al's (2001) study, reliability estimates for six of the seven subscales were above .70 except for the autonomy subscale (a =.57). In the present study the reliability estimates were as follows: Change (a = .64), Cohesion (a = .87), Communication (a = .84), Mission (a = .70), Stress (a = .81), Autonomy (a = .39), and Total Score (a =.84). See Table 2 for supervision group level descriptive statistics for this measure.

Youth Treatment Engagement. Youth clients (n = 910) receiving services from study therapists completed the My Thoughts About Therapy Questionnaire (MTT; Chorpita & Becker, 2022) after completing at least four sessions of therapy. The MTT is a 35-item, self-report measure that assesses five domains of treatment engagement as identified by Becker et al. (2018): Relationship (i.e., therapeutic alliance; e.g., "I feel like I can tell my counselor anything"), Expectancy (i.e., expectation of positive treatment outcomes; e.g., "I believe the work I do with my counselor will help me"), Attendance (i.e., ability to attend treatment as scheduled; e.g., "I make sure I get to my appointments with my counselor"), Clarity (i.e., understanding of treatment rationale, goals and role of each participant: e.g., "I understand what I'm supposed to be doing in counseling"), and Homework (i.e., completion of homework and active participation during in-session activities, "I enjoy practicing new things with my counselor"), or REACH. Each item was scored on a four-point scale from 0 (strongly disagree) to 3 (strongly agree), with lower scores indicative of poorer engagement in that domain. Consistent with Chorpita and Becker (2022), records were removed listwise if any of the five subscales were missing at least three items or if the full measure was missing at least seven items. Otherwise, missing values were computed using person mean substitution which has been shown to be an appropriate method for estimating missing values in simulation studies when

both the number of respondents with missing data and number of items missing for each subscale is less than 20% though some researchers have suggested this approach is appropriate for up to 50% of items missing on a given scale (Downey & King, 1998; Hawthorne & Elliott, 2005). To be eligible for participation in the parent Reaching Families Study, youth had to have a score on the MTT that was indicative of possible engagement difficulties. However, for the purpose of these analyses, all youth who completed the MTT has part of routine LAUSD screening and who were being seen by a study therapy were included.

Candidate Predictors of Climate

Attitudes. Supervisors completed the Evidence-Base Practice Attitudes Scale 50-item version (EBPAS-50; Aarons et al., 2012). For the purposes of the current study, the 15 items associated with the original, highly utilized EBPAS scale were examined (Aarons, 2004) as the total score from the EBPAS-15 is commonly used to represent global attitudes towards EBPs and previous studies have established United States National norms for the EBPAS-15 thus adding helpful context for scores observed in the current study (Aarons et al., 2010) Each item was scored on a five-point scale from 0 (not at all) to 4 (to a very great extent), with lower scores indicative of worse attitudes towards EBPs. The total score was calculated using the mean of the 15 items. In Aaron's (2004) study, the reliability estimate for the total scale was a = .77. In the current study, the overall scale reliability estimate was a = .75. Records were removed if more than three items (20%) were missing. Missing values were computed using person mean substitution (Downey & King, 1998; Hawthorne & Elliott, 2005). In the current dataset, 22.7% (n = 5) respondents had any missing data and 13.7% of participants (n = 3) had more than three missing items. Supervisor EBPAS-15 scores ranged from 2.47 to 3.80 with a mean of 3.25 (SD = .42)

Communication. In addition to reflecting on communication at the organizational level via the TCU-ORG, therapists were also asked to reflect on communication at supervisory team level via the same set of Communication subscale items with different instructions: "Please consider your supervision team when answering the following questions again." Referent-shift consensus items refer to items that are worded such to prompt respondents to answer questions in relation to the same level of aggregation, in this specific case to answer questions in relation to the supervision team to which they belong (Chan 1998). Missing data for this scale were handled the same way as missing data from the broader TCU-ORG scale; 3.5% of therapists (n = 3) were missing more than half of the items on the subscale and the score could not be calculated.

Caseload. In the background questionnaire, supervisors were asked to report their current supervisory caseload: "How many active supervision cases do you typically carry at one time?" Supervisors reported between 0 and 150 supervision cases (M = 36.83, SD = 48.58). Examination of a histogram of this item revealed that two supervisors reported caseloads of between 50 and 100, two supervisors reported caseload sizes between 100 and 150, and the rest reported caseload sizes less than 50. Examination of the boxplot revealed that values of 130 and 150 were outliers in the data. In addition, four supervisors did not answer this question. The wide variability of the data raises questions about the clarity of the item wording and validity of responses. It is possible that some supervisors were answering in response to the *number of providers* they supervise while other supervisors were answering in response to the *number of clients* seen by the providers they supervise.

Similar issues were seen in response to the question about supervisor ideal caseload: "How many cases would you say constitute an appropriate supervision caseload – one that would allow you to do your best work with each supervisee?" Supervisors reported between 2 and 80

supervision cases (M = 21.63, SD = 21.87). Examination of the histogram of this item revealed that four supervisors reported ideal caseloads of between 40 and 80, while the majority reported ideal caseload sizes of less than 20. Examination of the boxplot revealed that the value 80 was an outlier in the data. In addition, three supervisors did not answer this question. Given issues with both items querying supervisory caseload and ideal caseload, caseload and discrepancy between caseload and ideal caseload were not used as predictors in Aim 2 as originally planned.

Tenure in the Organization. In the background questionnaire, supervisors were asked about their tenure in the organization: "How many years have you worked at [site]?" Supervisors reported between 1 and 20 years (M = 12.64, SD = 5.33). However, given an error in data collection, there was missing data for eight supervisors (36.4% of the supervisor sample). For this reason, tenure was not used as a predictor in Aim 2 as originally planned.

Data Analysis

Aim 1. The first aim of the study was to determine the extent to which therapists' perceptions of organizational climate was shared within supervision team and unique from other supervision teams. Agreement indices capture the extent to which scores generated by different respondents are equivalent in terms of absolute value (e.g., both Therapist A and Therapist B from Supervision Group 1 rate climate as a "3" on a five-point likert scale; Bliese et al., 2000). Although historically the $r_{wg(j)}$ index has been used to characterize agreement for multi-item scales, it has several limitations (Brown and Hauenstein, 2005). First, $r_{wg(j)}$ uses a uniform null distribution (i.e., assumes all likert scale points are equally likely to be selected) for comparison when estimating agreement. This is problematic because it does not account for rater bias (e.g., tending to use the higher end of the scale because of social desirability; Bliese, 2000). In addition, the $r_{wg(j)}$ index is influenced by sample size and number of points on the likert scale

(Brown & Hauenstein, 2005). Some researchers have recommended use of other null response distributions (e.g., Beimann et al., 2012) to overcome the first limitation. However, Brown and Hauenstein (2005) argue that it is challenging to determine which null distribution is valid for a given scale and instead propose use of the $a_{wg,}$ which addresses the above limitations. In the current study, both the $r_{wg(j)}$ and the a_{wg}, were calculated, enabling readers to compare this study's findings with previous literature ($r_{wg(j)}$) while also generating the most accurate index of agreement per best-practice recommendations from the organizational management literature (a_{wg}).

Although agreement indices are useful for determining equivalence of ratings given by therapists in the same supervision group, they do not provide information about the extent to which therapists from different supervision groups are distinct from each other. This means that a high value observed for a_{wg} could be the result of high agreement within individual groups or high agreement across the entire sample. For this reason, it is important to contextualize agreement indices by pairing them with reliability indices. Reliability indices capture relative consistency of responses among respondents (e.g., In Supervision Group 1, Therapist A rates items one through three as "2", "3", "4" and Therapist B rates items one through three as "3", "4", and "5"; Bliese et al., 2000). Two types of interclass correlation coefficients (ICC (1) and ICC (2)) were calculated using multilevel models (therapists nested within supervisors) with random intercepts and no predictors for each climate subscale (i.e., Mission, Cohesion, Autonomy, Stress, and Change) and total score. ICC (1) is a measure of the proportion of total variance explained by group membership (Bryk and Raudenbush, 1982) and was calculated by dividing between-group variance by total variance. A likelihood ratio test, which compares a model with random intercepts and no predictors to a model with no random intercepts and no

predictors was used to determine if ICC (1) values were significantly different from zero. ICC (2) is a measure of the reliability of group means (e.g., mean ratings reliably distinguish groups; Le Breton & Senter, 2008) and was calculated by dividing the difference of between and within group variance by between group variances. A lower ICC (2) can occur when members across groups provide similar ratings.

Aim 2. The second aim of the study was to investigate the association between candidate supervisor characteristics and therapist perception of climate subscales and climate total score. Given that therapists were nested within supervisors, multilevel models with random intercepts were used to assess the association between supervisor attitudes towards EBPs and supervisor communication on provider perception of climate. School district (LAUSD, CA; Santee Wateree, SC; Pee Dee, SC) was included as a covariate given known differences in context including frequency and amount of time spent in supervision, case load sizes, and policies supporting EBT use (i.e., Los Angeles County Department of Mental Health Prevention and Early Intervention Plan, 2009). Assumptions of multilevel modeling were met, including normality of residuals for both level 1 and level 2 predictors, homoscedasticity, and assumptions of linearity.

Aim 3. The third aim of the study was to determine if organizational climate scores aggregated at supervisory team level (i.e., mean score for each team) were predictive of client treatment engagement as measured by the youth MTT. Clients were nested within therapists who were nested within supervisor teams, therefore multilevel modeling was used to assess the impact of organizational climate subscales and total score aggregated at the supervisory level on client engagement. Given potential differences in climate perceptions from individual therapists on the same supervision teams, climate strength (i.e., variability of perceptions as measured by standard deviation; Schneider et al., 2013) was tested as a moderator between supervisor team-

level climate and client engagement using procedures outlined by Hayes (2017). Assumptions of multilevel modeling including normality of residuals, homoscedasticity, and assumptions of linearity were met. All statistical analyses were carried out using R 4.1.2 (R Core Team, 2021) *multilevel* (version 2.6; Bliese, 2016) *nlme* (version 3.1-155; Pinheiro et al., 2022), and *psych* (version 2.2.3; Revelle, 2022) packages.

Results

Aim 1

The first study aim was to evaluate the extent to which therapists' perceptions of organizational climate were shared within supervision team and unique from other supervision teams using two agreement indices ($r_{wg(j)}$, and a_{wg}) and two reliability indices (ICC (1) and ICC (2)). Values ranged from mean $r_{wg(j)}$ .61 for Cohesion to mean rating $r_{wg(j)}$ .74 for Autonomy and Change. Values ranged from mean a_{wg} .61 for Cohesion to mean a_{wg} .74 for Autonomy (see Table 2 for mean and range and Table 3 for a_{wg} indices for each supervision group). Table 2 includes the ICC (1) and ICC (2) for each TCU-ORG subscale and total score. ICC (1) values ranged from .00 for the subscale, Mission to .38 for the subscale Autonomy. Likelihood ratio tests indicated that ICC (1) values for Autonomy (ICC (1) = .38) and Cohesion (ICC (1) = .19) were significantly different from zero. Because the ICC (1) was so low for the Mission subscale, the TCU total score and ICC (1) was recalculated without it. This resulted in a new ICC (1) of .18 which was significantly different than zero as indicated by the likelihood ratio test. ICC (2) values ranged from .00 for the subscale Mission to .66 for the subscale Autonomy. The ICC (2) value for Cohesion was .44. The collective findings show that the therapists on the same supervision team shared perceptions of organizational climate that were similar to each other and distinct from therapists on different supervision teams for the subscales Autonomy, Cohesion, and TCU total score with Mission removed.

Aim 2

The second aim of the study was to investigate the association between candidate supervisory characteristics (supervisor attitudes towards EBPs measured by the supervisor self-report EBPAS and therapist perception of supervisor communication) and therapist perceptions of climate subscales. Cohesion, Autonomy, and total climate score (without Mission) were the three scales for which a significant amount of variance was attributable to supervision group membership. Because the present study centers on supervisors' influence on shared group perceptions of climate, only these scales were used as outcome variables in Aim 2.

Autonomy. Autonomy was regressed on school district (Pee Dee SC; Santee Wateree, SC; LAUSD, CA), supervisor attitudes toward EBPs, and therapist rated supervisor communication in a multi-level model with therapists nested within supervisors. In this model, school district was a significant predictor of Autonomy such that therapists at Pee Dee and Santee Wateree rated Autonomy significantly higher than therapists at LAUSD (Pee Dee $b = .35$, $p < .01$; Santee Wateree $b = .25$, $p < .05$), over and above the effect of supervisor attitudes towards EBPs and therapist rated supervisor communication. Therapist rated supervisor communication was a significant predictor of Autonomy such that for every one unit increase in therapist rated supervisor communication, Autonomy ratings increased by .34 ($b = .34$, $p < .01$) over and above the effect of supervisor attitudes towards EBPs and school district. In other words, therapists perceived more freedom and latitude in doing their jobs when supervisors engaged in clear, bidirectional communication that kept therapists informed. Supervisor attitudes towards EBPs was not a significant predictor of Autonomy ($b = -.04$, $p = .74$; see Table 4).

Cohesion. Cohesion was regressed on school district (Pee Dee SC; Santee Wateree, SC; LAUSD, CA), supervisor attitudes toward EBPs, and therapist rated supervisor communication in a multi-level model with therapists nested within supervisors. In this model, therapist rated

supervisor communication was a significant predictor of Cohesion such that for every one unit increase in therapist rated supervisor communication, Cohesion ratings increased by .70 (b = .70, p < .01) over and above the effect of supervisor attitudes towards EBPs and school district. Therapists perceived higher workgroup cooperation and teamwork when supervisor communication was strong. Supervisor attitudes towards EBPs (b = .15, p = .51) and school district (Pee Dee b = .06, p = .80; Santee Wateree b = .11, p = .61) were not significant predictors of Cohesion (see Table 5).

Total. Total climate score excluding the Mission subscale was regressed on school district (Pee Dee SC; Santee Wateree, SC; LAUSD, CA), supervisor attitudes toward EBPs and therapist rated supervisor communication in a multi-level model with therapists nested within supervisors. In this model, therapist rated supervisor communication was a significant predictor of Total Climate score such that for every one unit increase in therapist rated supervisor communication, Total Climate score ratings increased by .59 (b = .59, p < .01) over and above the effect of supervisor attitudes towards EBPs and school district, suggesting that therapist perception of supervisor communication had generalized positive effects on perception of climate overall. Therapists in the school district, Pee Dee, SC, had significantly higher rating of climate compared with LAUSD therapists (b = .21, p < .05). Therapists in the school district, Santee-Wateree, SC, did not (b = .16, p = .08). Supervisor attitudes towards EBPs (b = .15, p = .51) was not a significant predictor of Total climate score (see Table 6).

Aim 3

The third aim of the study was to determine if organizational climate scores aggregated at supervisory team level (i.e., mean score for each team) were predictive of client treatment engagement and moderated by climate strength (i.e., standard deviation for each team). Given

that Cohesion, Autonomy, and the total score (without the Mission subscale) were the scales for which a significant amount of variance was attributable to group membership, these scales were used as predictor variables in three separate models. Clients were nested within therapists who were nested within supervisor teams. School district was included as a covariate in each model. Autonomy (b = 15.79, p = .17), Cohesion (b = 13.09, p = .34) and total score (b = 8.28, p = .45) were not predictive of youth engagement nor was climate strength a moderator in any of the three models. Further, both supervisor level and therapist level did not account for a significant amount of variance in youth engagement as indicated by a likelihood ratio test comparing a null model to an intercept only model (supervisor level ICC (1) = .01; therapist level ICC (1) =.02) (see Tables 7 through 9 for each model).

Discussion

The main goal of the present study was to explore the role of clinical supervisors as drivers of therapist perceptions of organizational climate and subsequent impact on client treatment engagement. Specifically, this study explored perception of organizational climate among therapists nested in supervision teams, investigated candidate predictors of therapists' perception of climate, and assessed the association between supervision team climate and client treatment engagement.

Aim 1: Perceptions of Organizational Climate

Within in the psychology literature, organizational constructs (e.g., organizational climate) tend to focus on individual provider perceptions rather than considering organizational sub-units (for exception see work from Glisson and colleagues such as Glisson et al., 2002), potentially missing useful insight that could be gained through examination of various organizational levels. However, in the broader organization management and leadership

literature there is a longer history of exploring how organizational units (e.g., hospitals, army bases, warehouses) and sub-units (e.g., clinical teams, platoons, shift managers) impact critical organizational outcomes (e.g., turnover, ethical behavior, accuracy in medication administration; Chan, 2014; Zohar & Luria, 2005). In the organizational management literature, there is also extensive discussion of indices that provide evidence of meaningful organizational subunits. Guided by this literature, Aim 1 evaluated the extent to which therapists' perceptions of organizational climate were shared within supervision teams and unique from other supervision teams using several indices including $a_{wg,}$ and $r_{wg(j)}$, ICC (1), and ICC (2). Interpretation of each index will be reviewed followed by discussion of the larger picture painted by indices together.

Agreement indices capture the extent to which scores generated by different therapists are equivalent in terms of absolute value (Bliese et al., 2000). For this reason, they are particularly useful for determining the appropriateness of aggregating individual scores to form a composite group score. Given limitations associated with the $r_{wg(j)}$ index described in the Methods section, Table 2 contains both $r_{wg(j)}$ and a_{wg} indices so the reader can see the differing results for each approach. However, only the a_{wg} results will be discussed herein.

An a_{wg} index was generated for each supervision group (see Table 2 for descriptive statistics and Table 3 for the a_{wg} index for each group). Mean a_{wg} for each subscale ranged from .61 for Cohesion, representing strong agreement to .74 for Autonomy, representing very strong agreement (LeBreton & Senter 2008), For each subscale, there was variability in agreement across groups with some groups having nearly perfect agreement and others having almost no agreement. One supervision group comprised of four therapists (labeled as Group 19 in Table 3) had highly conflicting ratings across all TCU-ORG subscales. Differences in agreement between groups about organizational climate speak to the concept of climate strength. Researchers such as

Schneider et al. (2012) have suggested that organizations in which individuals have markedly different experiences of climate may be distinct from those in which individuals tend to agree. Disagreement could be an indication of a breakdown in communication, individuals feeling alienated from the larger group, or differential treatment by management. Climate strength will be explored further in the discussion of Aim 3.

ICC (1) is a measure of the proportion of total variance explained by group membership (Bryk and Raudenbush, 1982) and can be expressed as a percentage. Autonomy and Cohesion were the two scales for which group membership explained a significant amount of variance as indicated by the likelihood ratio test. Specifically, group membership explained approximately 38% of the variance in Autonomy ratings and 19% of variance in Cohesion ratings representing large and medium effects respectively. Notably, group membership explained almost no variance in Mission ratings (ICC (1) = .00). Because the ICC (1) was so low for the Mission subscale, the total score and ICC (1) was recalculated without it. This resulted in a new ICC (1) of 18% which was significantly greater than zero as indicated by the likelihood ratio test. ICC (2) is a measure of the reliability of group means (Le Breton & Senter, 2008). In the current study, supervision group sizes ranged from 1 to 7 therapists per group with a mean of 3.91 therapists per group. Mission had the lowest (ICC (2) = .00) and Autonomy had the highest (ICC (2) = (.66)). However, ICC (2) is highly dependent on group size, and it is challenging to achieve values of .70 in smaller groups (5-6 individuals per group; Schneider et al., 2013).

Taken together, these indices paint a more complete picture of how supervision groups influence perceptions of climate. Consistent with the first hypothesis (H1) that therapists in the same supervisory team will have more shared perceptions of organizational climate compared with therapists in different supervisory teams, team membership accounted for a significant

amount of variance for Autonomy, Cohesion, and total climate score (removing the subscale, Mission). Interestingly, results also suggest that subdomains of organizational climate may be differentially influenced by group membership. Taking the most extreme examples, supervision group explained almost no variance in therapists' perception of Mission. However, agreement about Mission for therapists across all groups was generally strong. Conversely, supervision group explained a large, statistically significant amount of variance in Autonomy and Cohesion. Therapists in the same group also tended to agree about organizational Autonomy and Cohesion and mean ratings for each group were somewhat reliable when distinguishing between groups. This pattern might suggest that although therapists, across all supervisory groups, had a shared understanding of organizational mission, their day-to-day interpersonal relationships with other employees (i.e., cohesion) and independence to conduct their work (i.e., autonomy) was influenced by their direct supervisor.

This pattern maps onto Birken et al.'s (2012) theory on the role of middle managers in the implementation of innovations. Specifically, organizational mission and priorities are expressed at the upper levels of the organization and would therefore be expected to remain stable regardless of supervision group. In turn, middle managers translate the mission and priorities into day-to day interactions and activities on their teams thus impacting a sense of cohesion and autonomy (Bunger et al., 2019). It is especially notable that supervision team was associated with almost 40% of the variability in perceptions of autonomy. Although there are some conflicting findings in the literature, generally, autonomy has been associated with an increased sense of personal achievement and reduced emotional exhaustion at work, both important factors for preventing professional burnout (Yang & Hayes, 2020).

Supervision has been described as serving three key functions, one of which is characterized as restorative, i.e., support of supervisee emotional wellbeing (Milne et al., 2007; Proctor, 1987). With this in mind, it is surprising that supervision group was not associated with significant variability in therapist ratings of stress. It is possible that the influence of organizational factors outside of the control of the supervisor such as productivity standards and case load size, were so prominent as to negate any possible impact of the supervisor (Franco et al., 2016; Quinn et al., 2019). It may also be that the focus of supervision tended not to include restorative activities. The latter explanation is consistent with previous literature examining activities in supervision-as-usual in the context of community mental health serving youth. In their study, Bailin et al.(2018) found the most common supervision activities were administrative tasks and praise (both occurring in almost 90% of supervision sessions coded for about a quarter of the time in supervision). However, a specific focus on supervisee wellbeing occurred in less than half of sessions and tended to take up less than 5% of the time in supervision. It is important to acknowledge that it is possible that other activities in supervision decrease therapist stress beyond an explicit focus on wellbeing.

Aim 2: Predictors of Climate

Supervisor attitudes toward EBPs was not a predictor of therapist perception of overall climate, cohesion, or autonomy. This finding was not consistent with the study's second hypothesis (H2.1) that supervisor would be significantly and positively associated with therapist perceptions of organizational climate. This is surprising given past research that found an association between therapist attitudes and aspects of organizational climate (Aarons & Sawitzky, 2006; Aarons & Sommerfeld, 2012). Based on these previous findings, one might anticipate that, because supervisors have greater organizational power and influence, supervisor

attitudes would serve as a powerful driver of organizational climate. However, attitudes toward EBPs from supervisors in the current sample were relatively higher and more homogenous than national norms (present study: $M = 3.25$, $SD = .42$; national norms: $M = 2.73$, $SD = .49$; Aarons et al., 2010). It is therefore possible that a sample with greater variability and more negative views of EBPs may have yielded different results.

Consistent with the study's third hypothesis (H2.2), therapist perception of supervisor communication was found to be significantly associated with overall climate, cohesion, and autonomy. It is possible therapists were aware of the overarching organizational mission but their direct work experience (cohesion among employees and autonomy to conduct their work) was influenced by the direct communication of the supervisor. This finding, again, aligns with emerging evidence that middle managers shape their employees' experience through communication strategies (Birken et al., 2012; Bunger et al., 2019).

Aim 3: Climate as a Predictor of Youth Engagement

The third aim of the study was to determine if supervision team shared perception of organizational climate was a predictor of youth treatment engagement and moderated by climate strength. The findings did not support hypothesis (H3). Supervision team climate was not a predictor of youth treatment engagement nor was the relation between climate and engagement moderated by climate strength. This is inconsistent with previous research linking client engagement with climate (Greener et al., 2006; Lehman et al., 2002). In the current study, supervisor and therapist level explained very little variance in client engagement (supervisor-level ICC (1) = .01; therapist-level ICC (1) = .02). This finding would suggest that variables at the broader system level, therapist-client dyad level or individual client-level are drivers of treatment engagement rather than specific characteristics of the therapist or supervisor. This

aligns with Becker et al.'s (2018) conceptualization of treatment engagement as a "multidimensional commitment to treatment" and a "dynamic process" that involves the reciprocal influence of individuals, families, professionals, the service organization, and community.

Limitations and Future Directions

This study had several limitations. First, the level 2 (supervisor) sample size was small thus increasing the chances of type 1 error for analyses using supervisor level predictors (McNeish & Stapleton, 2014). Though it is possible that a larger sample of the supervisor population may have yielded different results, in simulation studies, level-two samples as low as 15 yielded unbiased estimates of fixed effects. Conversely, level-two samples of 30 or higher may be necessary for unbiased estimates of fixed-effect standard errors. Samples smaller than 30, therefore, increase the chances of type-I error (McNeish & Stapleton, 2014). However, despite the potential for inflated *p*-values, supervisor attitudes of EBTs were not found to be a predictor of therapist perception of climate, perhaps increasing confidence in this result.

Second, therapist perception of supervisor communication was collected concurrent with perceptions of broader organizational climate. The association between supervisor communication and climate scales, may therefore, be a result of a response bias (e.g., tendency to respond positively across all scales due to a halo effect). Future research might consider exploring this association further by using objective measures of supervisor communication strategies. However, a therapist's impression of how their supervisor communicates (rather than the "true" quality of communication) may be uniquely important for informing how therapists feel about the organization they work in. Third, organizations are often comprised of many units and subunits, and it is challenging to know which units are impacting an employees' experience

of their work environment (e.g., care network, hospitals, department, unit, shift). Although this study dealt specifically with the supervision level and accounted for district level through inclusion as a covariate, other levels of possible importance were not assessed. For example, in addition to the school district level, therapists (n = 45) and their supervisors (n = 12) at LAUSD tend to work in specific "neighborhoods" comprised of several schools. The districts in South Carolina (Pee Dee and Santee Wateree) were much smaller and therapists (n = 41) and supervisors (n = 10) tend to move between multiple schools in the district making the influence of a school or neighborhood level less likely. However, given that LAUSD therapists comprised approximately half of the sample the grouping seen in Aim 1 may have been driven by some feature of the neighborhood instead of or in addition to the characteristics of the supervisor.

Third, providers were excluded from the present study if they were in supervisory dyads that were created for the purpose of the study. It is possible that the exclusion of these providers may have meant that providers with the most demanding schedules, that necessitated creation of supervisory dyads, were not represented. Results should be taken with this caveat in mind. However, it was important to the study design that providers were clearly nested within supervisory teams so this trade off was necessary.

Finally, in the present study, client engagement was measured by taking the full score of the MTT. However, it is possible that certain aspects of treatment engagement may be differentially impacted by organizational climate. The overall engagement score was used in the current study to align with past studies connecting engagement with climate (Lehman et al., 2001; Greener et al., 2007). However, future studies might explore the association between climate and components of treatment engagement (e.g., relationship between therapist and client,

expectations about treatment) to identify possible differential impact on specific aspects of engagement.

The findings that therapist perception of the communication on their supervision team was associated with perception of organizational climate overall, cohesion, and autonomy provide evidence to suggest that supervisors play an important role in shaping perceptions of climate. To better disentangle the role of the supervisor from aspects of the organization at other levels, future research should further explore the association between climate and supervisory level characteristics such as leadership style (Aarons & Sommerfeld, 2012), working alliance, and supervisory behaviors (Dorsey et al., 2018). Future studies might also identify and test strategies for increasing the ability of the supervisor to have an impact on organizational climate such as examining the impact of additional protected time for supervision or training in specific supervisory behaviors or activities that foster a positive perception of the organization.

Conclusions

The present study capitalized on prior research and methodology from the organizational management literature to explore the role of supervisors in shaping therapists experience of their work environment. Results build upon past research describing the key role of middle managers as intermediaries between upper-level management and front-line workers, particularly in terms of perceptions of cohesion and autonomy. Further elucidating the role of supervisors in mental health organizations will be important for maximizing workforce capacity to address the substantial and growing demand for mental health treatment.

Tables

Table 1.

Demographics of Supervisor and Therapists

	Supervisor (n = 22)		Therapists (n = 86)		Total (n = 108)	
	N	%	*N*	%	*N*	%
Race/Ethnicity						
Asian	1	4.55	2	2.33	3	2.78
Black/African American	11	50.00	36	41.86	47	43.52
White/Caucasian	4	18.18	13	15.12	17	15.74
Hispanic/Latino	4	18.18	34	39.53	38	35.19
Multiracial	2	9.09	1	1.16	3	2.78
Gender						
Woman	21	95.45	78	90.70	99	91.67
Man	1	4.55	8	9.30	9	8.33
School District						
LAUSD	12	54.55	45	52.33	57	52.78
Santee Wateree	6	27.27	22	25.58	28	25.93
Pee Dee	4	18.18	19	22.09	23	21.30
Academic Degree						
Bachelor's-level	0	0.00	1	1.16	1	0.94
Master's-level	20	100.00	84	97.67	104	98.11
PhD	0	0.00	1	1.16	1	0.94
	M	*SD*	*M*	*SD*	*M*	*SD*
Age	47.15	10.53	38.60	10.15	40.42	10.77
Years of Clinical Experience	15.83	8.09	6.73	5.85	8.44	7.24

Note. Two supervisors did not provide years of experience or degree type. Eight providers did not provide age

Table 2.

Supervision Group Descriptive Statistics, Intraclass Correlation Coefficients, and Reliability Estimates for Climate Subscales and Total Score

Scale Name	*M (SD)*	Range	ICC (1)	ICC (2)	$r_{wg(j)}$ *M*	$r_{wg(j)}$ Range	a_{wg} *M*	a_{wg} Range
Autonomy	2.68 (.25)	2.16 to 3.00	**0.38**	0.66	0.90	.30 to .98	0.74	.22 to .92
Change	2.73 (.24)	2.20 to 3.07	0.10	0.28	0.88	.00 to .99	0.73	.08 to .96
Cohesion	2.92 (.35)	2.17 to 3.50	**0.19**	0.44	0.86	.00 to .99	0.61	-.38 to .96
Communication	2.54 (.37)	1.72 to 3.15	0.19	0.44	0.86	.00 to .98	0.68	.01 to .92
Mission	3.01 (.18)	2.72 to 3.40	0.00	0.00	0.92	.51 to .99	0.73	-.75 to 1.00
Stress	2.75 (.35)	2.19 to 3.50	0.17	0.42	0.83	.00 to .97	0.64	.13 to 1.00
Total	2.78 (.18)	2.38 to 3.07	0.16	0.39	0.93	.00 to 1.00	0.67	-.06 to .90
Total without Mission	2.65 (.25)	1.99 to 3.13	**0.18**	0.44	0.93	.00 to 1.00	0.67	-.03 to .90

Note. **Bold** indicates variance explained by level 2 (supervisors) was significantly different from 0 using likelihood ratio test, $p < .05$

Table 3.

Agreement Index (a_{wg}) by Supervision Group

Supervision Group	Therapists (*n*)	Autonomy	Cohesion	Communication	Change	Stress	Mission	Total	Total without Mission
1	5	0.72	0.39	0.76	0.89	0.76	0.71	0.68	0.68
2	6	0.85	0.65	0.82	0.58	0.60	0.83	0.73	0.71
3	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4	4	0.92	0.91	0.65	0.77	0.52	0.76	0.76	0.76
5	4	0.71	0.68	0.76	0.67	0.51	0.81	0.70	0.66
6	5	0.36	0.60	0.63	0.76	0.60	0.73	0.57	0.55
7	2	0.78	0.78	0.78	0.78	0.90	0.78	0.78	0.78
8	3	0.78	0.93	0.88	0.96	0.85	0.88	0.78	0.78
9	4	0.76	0.93	0.53	0.92	0.79	0.81	0.79	0.80
10	2	0.78	0.78	0.89	0.78	1.00	0.78	0.82	0.83
11	5	0.91	0.55	0.72	0.73	0.83	0.86	0.75	0.73
12	4	0.90	N/A	0.78	0.80	0.67	1.00	0.83	0.79
13	6	0.81	0.46	0.75	0.69	0.68	0.81	0.70	0.68
14	3	0.73	0.30	0.63	0.70	0.33	0.95	0.65	0.59
15	6	0.86	0.55	0.62	0.71	0.63	0.74	0.64	0.62

16	3	0.80	0.96	0.92	0.95	0.80	0.95	0.90	0.90
17	6	0.69	0.70	0.59	0.67	0.42	0.59	0.62	0.63
18	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
19	4	0.22	-0.38	0.01	0.08	0.13	-0.75	-0.06	-0.03
20	7	0.68	0.60	0.51	0.69	0.45	0.83	0.66	0.62
21	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
22	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note. a_{wg} index cannot be calculated for groups of one

Table 4.

Predictors of Therapist Perception of Autonomy in Organization

Factor	*b*	*SE*	*p*
Intercept	1.81	0.37	0.00
EBPAS-15	-0.04	0.11	0.74
Communication	0.34	0.05	0.00
District			
Pee Dee	0.35	0.10	0.01
Santee Wateree	0.25	0.11	0.03
Variance Components	Variance	*SD*	
Residual	0.04	0.20	
Intercept	0.02	0.13	

Note. Therapists (*n* = 76), supervisors (*n* = 19); LAUSD was the reference group for District

Table 5.

Predictors of Therapist Perception of Autonomy in Organization

Factor	*b*	*SE*	*p*
Intercept	0.57	0.73	0.44
EBPAS-15	0.15	0.22	0.51
Communication	0.70	0.10	0.00
District			
Pee Dee	0.06	0.22	0.80
Santee Wateree	0.11	0.21	0.61
Variance Components	Variance	*SD*	
Residual	0.17	0.41	
Intercept	0.06	0.25	

Note. Therapists (n = 76), supervisors (n = 19); LAUSD was the reference group for District

Table 6.

Predictors of Therapist Perception of Organizational Climate

Factor	*b*	*SE*	*p*
Intercept	0.93	0.30	0.00
EBPAS-15	0.03	0.09	0.73
Communication	0.59	0.06	0.00
District			
Pee Dee	0.21	0.09	0.03
Santee Wateree	0.16	0.09	0.08
Variance Components	Variance	*SD*	
Residual	0.06	0.03	
Intercept	0.00	0.24	

Note. Therapists (n = 76), supervisors (n = 19); LAUSD was the reference group for District

Table 7.

Supervision Team Autonomy and Climate Strength Predicting Youth Baseline Treatment Engagement

Factor	*b*	*SE*	*p*
Intercept	33.29	29.78	0.26
Autonomy	15.79	10.94	0.17
Strength	165.11	91.13	0.09
Autonomy x Strength	-63.09	33.85	0.09
District			
Pee Dee	-2.44	2.90	0.42
Santee Wateree	-1.20	2.97	0.69
Variance Components	Variance	*SD*	
Residual	362.05	19.03	
Supervisor (Intercept)	2.93	1.71	
Therapist (Intercept)	3.68	1.92	

Note. Only supervision groups with more than one therapist were used in this model; students (n = 910), therapists (n = 82), supervisors (n = 19); LAUSD was the reference group for District

Table 8.

Supervision Team Cohesion and Climate Strength Predicting Youth Baseline Treatment Engagement

Factor	*b*	*SE*	*p*
Intercept	35.34	38.78	0.36
Cohesion	13.09	13.17	0.34
Strength	92.30	82.55	0.28
Cohesion x Strength	-30.51	28.32	0.30
District			
Pee Dee	-3.63	2.25	0.13
Santee Wateree	-1.87	2.27	0.42
Variance Components	Variance	*SD*	
Residual	361.18	19.00	
Supervisor (Intercept)	3.21	1.79	
Therapist (Intercept)	5.27	2.30	

Note. Only supervision groups with more than one therapist were used in this model; Students (n = 910), therapists (n = 82), supervisors (n = 19); LAUSD was the reference group for District

Table 9.

Supervision Team Climate Total Score and Climate Strength Predicting Youth Baseline Treatment Engagement

Factor	*b*	*SE*	*p*
Intercept	52.52	27.78	0.06
Total	8.28	10.56	0.45
Strength	129.83	117.84	0.29
Total x Strength	-47.79	44.29	-1.08
District			
Pee Dee	-2.52	2.38	-1.06
Santee Wateree	-0.88	2.63	-0.33
Variance Components	Variance	*SD*	
Residual	361.05	19.00	
Supervisor (Intercept)	3.52	1.88	
Therapist (Intercept)	5.43	2.33	

Note. Only supervision groups with more than one therapist were used in this model; Students (n = 910), therapists (n = 82), supervisors (n = 19); LAUSD was the reference group for District

References

Aarons, G. A., Glisson, C., Hoagwood, K., Kelleher, K., Landsverk, J., & Cafri, G. (2010). Psychometric properties and US National norms of the Evidence-Based Practice Attitude Scale (EBPAS). *Psychological Assessment*, *22*(2), 356.

Aarons, G. A., & Sawitzky, A. C. (2006). Organizational climate partially mediates the effect of culture on work attitudes and staff turnover in mental health services. *Administration and Policy in Mental Health and Mental Health Services Research*, *33*(3), 289. https://doi.org/10.1007/s10488-006-0039-1

Aarons, G. A., & Sommerfeld, D. H. (2012). Leadership, innovation climate, and attitudes toward evidence-based practice during a statewide implementation. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(4), 423-431. https://doi.org/10.1016/j.jaac.2012.01.018

Aarons, G. A., Sommerfeld, D. H., & Willging, C. E. (2011). The soft underbelly of system change: The role of leadership and organizational climate in turnover during statewide behavioral health reform. *Psychological Services*, *8*(4), 269. https://doi.org/10.1037/a0026196

Bailin, A., Bearman, S. K., & Sale, R. (2018). Clinical supervision of mental health professionals serving youth: Format and microskills. *Administration and Policy in Mental Health and Mental Health Services Research*, *45*(5), 800-812. https://doi.org/10.1007/s10488-018-0865-y

Becker, K. D., Boustani, M., Gellatly, R., & Chorpita, B. F. (2018). Forty years of engagement research in children's mental health services: Multidimensional measurement and

practice elements. *Journal of Clinical Child & Adolescent Psychology*, *47*(1), 1-23. https://doi.org/10.1080/15374416.2017.1326121

Biemann, T., Cole, M. S., & Voelpel, S. (2012). Within-group agreement: On the use (and misuse) of rWG and rWG (J) in leadership research and some best practice guidelines. *The Leadership Quarterly*, *23*(1), 66-80. https://doi.org/10.1016/j.leaqua.2011.11.006

Birken, S. A., Lee, S. Y. D., & Weiner, B. J. (2012). Uncovering middle managers' role in healthcare innovation implementation. *Implementation Science*, *7*(1), 1-12. https://doi.org/10.1186/1748-5908-7-28

Bliese, P. (2016). Multilevel: Multilevel Functions. R package version 2.6. https://CRAN.R-project.org/package=multilevel

Borntrager, C., Chorpita, B. F., Higa-McMillan, C. K., Daleiden, E. L., & Starace, N. (2013). Usual care for trauma-exposed youth: Are clinician-reported therapy techniques evidence-based? *Children and Youth Services Review*, *35*(1), 133 141. https://doi.org/10.1016/j.childyouth.2012.09.018

Brimhall, K. C., Fenwick, K., Farahnak, L. R., Hurlburt, M. S., Roesch, S. C., & Aarons, G. A. (2016). Leadership, organizational climate, and perceived burden of evidence-based practice in mental health services. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(5), 629-639. https://doi.org/10.1007/s10488-015-0670-9

Bunger, A. C., Birken, S. A., Hoffman, J.A., MacDowell, H., Choy-Brown, M., & Magier, E. (2019) Elucidating the influence of supervisors' roles on implementation climate. *Implementation Science*, *14*(1), 1-12. https://doi.org/10.1186/s13012-019-0939-6

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, *83*(2), 234.

Chan, D. (2014). Multilevel and aggregation issues in climate and culture research. In B. Schneider & K. M. Barbera (Eds.), *The Oxford Handbook of Organizational Climate and Culture* (pp. 484–495). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199860715.013.0025

Chorpita, B. F., & Becker, K. D. (2022). Dimensions of treatment engagement among youth and caregivers: Structural validity of the REACH framework. *Journal of Consulting and Clinical Psychology*. *90*(3), 258–271. https://doi.org/10.1037/ccp0000711

Chorpita, B. F., Daleiden, E. L., Ebesutani, C., Young, J., Becker, K. D., Nakamura, B. J., ... & Starace, N. (2011). Evidence-based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice*, *18*(2), 154-172. https://doi.org/10.1111/j.1468-2850.2011.01247.x

Chorpita, B. F., Korathu-Larson, P., Knowles, L. M., & Guan, K. (2014). Emergent life events and their impact on service delivery: Should we expect the unexpected? *Professional Psychology: Research and Practice*, *45*(5), 387. https://doi.org/10.1037/a0037746

Colquitt, J. A., Noe, R. A., & Jackson, C. L. (2002). Justice in teams: Antecedents and consequences of procedural justice climate. *Personnel Psychology*, *55*(1), 83-109. https://doi.org/10.1111/j.1744-6570.2002.tb00104.x

Dorsey, S., Pullmann, M. D., Kerns, S. E., Jungbluth, N., Meza, R., Thompson, K., & Berliner, L. (2017). The juggling act of supervision in community mental health: Implications for supporting evidence-based treatment. *Administration and Policy in Mental Health and*

Mental Health Services Research, *44*(6), 838-852. https://doi.org/10.1007/s10488-017-0796-z

Downey, R. G., & King, C. V. (1998). Missing data in Likert ratings: A comparison of replacement methods. *The Journal of General Psychology*, *125*(2), 175-191. https://doi.org/10.1080/00221309809595542

Franco, G. E. (2016). Productivity standards: Do they result in less productive and satisfied therapists? *The Psychologist-Manager Journal*, *19*(2), 91. https://doi.org/10.1037/mgr0000041

Garland, A. F., Brookman-Frazee, L., Hurlburt, M. S., Accurso, E. C., Zoffness, R. J., Haine-Schlagel, R., & Ganger W. (2010). Mental health care for children with disruptive behavior problems: A view inside therapists' offices. *Psychiatric Services*, *61*(8), 788-795. https://doi.org/10.1176/ps.2010.61.8.788

Glisson, C., Dukes, D., & Green, P. (2006). The effects of the ARC organizational intervention on caseworker turnover, climate, and culture in children's service systems. *Child Abuse & Neglect*, *30*(8), 855-880. https://doi.org/10.1016/j.chiabu.2005.12.010

Glisson, C., & Green, P. (2011). Organizational climate, services, and outcomes in child welfare systems. *Child Abuse & Neglect*, *35*(8), 582-591. https://doi.org/10.1016/j.chiabu.2011.04.009

Glisson, C., & James, L. R. (2002). The cross-level effects of culture and climate in human service teams. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *23*(6), 767-794. https://doi.org/10.1002/job.162

Glisson, C., & Schoenwald, S. K. (2005). The ARC organizational and community intervention strategy for implementing evidence-based children's mental health treatments. *Mental Health Services Research*, *7*(4), 243-259. https://doi.org/10.1007/s11020-005-7456-1

Glisson, C., Schoenwald, S. K., Kelleher, K., Landsverk, J., Hoagwood, K. E., Mayberg, S., ... & Research Network on Youth Mental Health. (2008). Therapist turnover and new program sustainability in mental health clinics as a function of organizational culture, climate, and service structure. *Administration and Policy in Mental Health and Mental Health Services Research*, *35*(1-2), 124-133. https://doi.org/10.1007/s10488-007-0152-9

Green, A. E., Albanese, B. J., Cafri, G., & Aarons, G. A. (2014). Leadership, organizational climate, and working alliance in a children's mental health service system. *Community Mental Health Journal*, *50*(7), 771-777. https://doi.org/10.1007/s10597-013-9668-5

Greener, J. M., Joe, G. W., Simpson, D. D., Rowan-Szal, G. A., & Lehman, W. E. (2007). Influence of organizational functioning on client engagement in treatment. *Journal of Substance Abuse Treatment*, *33*(2), 139-147. https://doi.org/10.1016/j.jsat.2006.12.025

Hawthorne, G. & Elliott, P. (2005). Imputing cross-sectional missing data: Comparison of common techniques. *Australian & New Zealand Journal of Psychiatry*, *39*(7), 583-590. https://doi.org/10.1080/j.1440-1614.2005.01630.x

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford publications.

Hoagwood, K., Burns, B. J., Kiser, L., Ringeisen, H., & Schoenwald, S. K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services*, *52*(9), 1179-1189. https://doi.org/10.1176/appi.ps.52.9.1179

Institute of Behavioral Research. (2005). *TCU-ORGSurvey of Organizational Functioning (TCU-ORGSOF)*. Fort Worth: Texas Christian University, Institute of Behavioral Research. Available at ibr.tcu.edu

James, L. R., Choi, C. C., Ko, C. H. E., McNeil, P. K., Minton, M. K., Wright, M. A., & Kim, K. I. (2008). Organizational and psychological climate: A review of theory and research. *European Journal of Work and Organizational Psychology*, *17*(1), 5-32. https://doi.org/10.1080/13594320701662550

Lau, A. S., Gonzalez, J. C., Barnett, M. L., Kim, J. J., Saifan, D., & Brookman-Frazee, L. (2018). Community therapist reports of client engagement challenges during the implementation of multiple EBPs in children's mental health. *Evidence-Based Practice in Child and Adolescent Mental Health*, *3*(3), 197-212. https://doi.org/10.1080/23794925.2018.1455545

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815-852. https://doi.org/10.1177/1094428106296642

Lehman, W. E., Greener, J. M., & Simpson, D. D. (2002). Assessing organizational readiness for change. *Journal of Substance Abuse Treatment*, *22*(4), 197-209. https://doi.org/10.1016/S0740-5472(02)00233-7

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review*, *28*(2), 295-314. https://doi.org/10.1007/s10648-014-9287-x

Morse, G., Salyers, M. P., Rollins, A. L., Monroe-DeVita, M., & Pfahler, C. (2012). Burnout in mental health services: A review of the problem and its remediation. *Administration and*

Policy in Mental Health and Mental Health Services Research, *39*(5), 341-352. https://doi.org/10.1007/s10488-011-0352-1

Olin, S. S., Williams, N., Pollock, M., Armusewicz, K., Kutash, K., Glisson, C., & Hoagwood, K. E. (2014). Quality indicators for family support services and their relationship to organizational social context. *Administration and Policy in Mental Health and Mental Health Services Research*, *41*(1), 43-54. https://doi.org/10.1007/s10488-013-0499-z

Ouellette, R. R., Goodman, A. C., Martinez-Pedraza, F., Moses, J. O., Cromer, K., Zhao, X., ... & Frazier, S. L. (2020). A systematic review of organizational and workforce interventions to improve the culture and climate of youth-service settings. *Administration and Policy in Mental Health and Mental Health Services Research*, *47*(5), 764-778. https://doi.org/10.1007/s10488-020-01037-y

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2022). Nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-155. https://CRAN.R-project.org/package=nlme

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Revelle, W. (2022) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych

Schneider, B., Ehrhart, M. G., & Macey, W. H. (2013). Organizational climate and culture. *Annual Review of Psychology*, *64*, 361-388. https://doi.org/10.1146/annurev-psych-113011-143809

Schneider, B., Salvaggio, A. N., & Subirats, M. (2002). Climate strength: a new direction for climate research. *Journal of applied psychology*, *87*(2), 220. https://doi.org/10.1037/0021-9010.87.2.220

Schoenwald, S. K., Chapman, J. E., Kelleher, K., Hoagwood, K. E., Landsverk, J., Stevens, J., Glisson, C., Rolls-Reutz, J. & Research Network on Youth Mental Health. (2008). A survey of the infrastructure for children's mental health services: Implications for the implementation of empirically supported treatments (ESTs). *Administration and Policy in Mental Health and Mental Health Services Research*, *35*(1-2), 84-97. https://doi.org/10.1007/s10488-007-0147-6

Southam-Gerow, M. A., Chorpita, B. F., Miller, L. M., & Gleacher, A. A. (2008). Are children with anxiety disorders privately referred to a university clinic like those referred from the public mental health system? *Administration and Policy in Mental Health and Mental Health Services Research*, *35*(3), 168–180. https://doi.org/10.1007/s10488-007-0154-7

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *American Psychologist*, *61*(7), 671–689. https://doi.org/10.1037/0003-066X.61.7.671

Williams, N. J., Frank, H. E., Frederick, L., Beidas, R. S., Mandell, D. S., Aarons, G. A., ... & Locke, J. (2019). Organizational culture and climate profiles: relationships with fidelity to three evidence-based practices for autism in elementary schools. *Implementation Science*, *14*(1), 1-14. https://doi.org/10.1186/s13012-019-0863-9

Woltmann, E. M., Whitley, R., McHugo, G. J., Brunette, M., Torrey, W. C., Coots, L., ... & Drake, R. E. (2008). The role of staff turnover in the implementation of evidence-based practices in mental health care. *Psychiatric Services*, *59*(7), 732-737.

Zohar, D., & Luria, G. (2005). A multilevel model of safety climate: cross-level relationships between organization and group-level climates. *Journal of Applied Psychology*, *90*(4), 616.

Chapter 2

Accountability in Supervision to Increase Therapist Use of Evidence in Treatment

Abstract

Clinical supervision has been touted as a natural opportunity for promoting therapist application of research evidence in community mental health services. However, only limited research points to the strategies or activities in supervision that are likely to capitalize on this opportunity. One theory suggests that supervision behaviors that promote a sense of accountability—i.e., a perceived expectation and responsibility to fulfill certain duties or obligations— may represent key strategies by which supervision can promote use of evidence. Informed by Grossman and Burke-Smalley (2018)'s model of accountability from the organizational management literature, the present study aimed to: (1) identify the presence of "accountability strategies" during clinical supervision; (2) investigate the association between accountability strategies and therapist use of a clinical procedure supported by evidence during the subsequent psychotherapy sessions. Transcripts of supervision sessions (n = 192) and subsequent treatment sessions (n = 192) were observationally coded for supervision accountability strategies and therapist follow through use of a clinical procedure supported by evidence identified during supervision in the subsequent treatment session (i.e., "expected practice"). Binary, multi-level logistic regression models were used to assess the association between accountability strategies during supervision and therapist use of an expected practice during the subsequent therapy session. Results indicated explicit *selection* of the practice the therapist will use in their next treatment session, *preparation* through activities such as reviewing manuals, role plays, and modeling, and *review* of how use of the practice went in the past, increased the odds of therapist use of that practice during the next treatment session. There is a general sentiment in the research literature and field of clinical psychology at large that supervision is important though research directly testing this assumption is nascent. These

findings are notable as they provide evidence to the growing body of research demonstrating that supervision can directly influence therapist behavior. Further, the present study delineates three concrete strategies that can be used to influence behavior.

Introduction

Given the notable discrepancy between effective interventions identified through research and treatment as it is delivered in community mental health settings, there is substantial interest in understanding factors that help or hinder the use of evidence in these settings with the purpose of uncovering strategies for improving the quality of care received by individuals and families seeking mental health services. Use of evidence refers to application of relevant knowledge identified through systematic research to the decision or context at hand (Park et al., 2020; Gray et al., 2015). Within the mental health field, use of evidence is sometimes conceptualized as therapist use of a manualized treatment (i.e., evidence is packaged in the form of a manual for use). A and is important as it results in better clinical outcomes compared with treatment as usual (Weisz et al., 2006; Weisz et al., 2013). Clinical supervision has been touted as a natural opportunity for promoting therapist application of research evidence in community mental health settings (e.g., Dorsey et al., 2013; Powell et al., 2015). Specifically, supervision has gained attention as it commonly occurs in community mental health settings and offers an opportunity for ongoing support for therapeutic skill development (Schoenwald et al., 2008; Bailin et al., 2018). This is key given the consistent finding that training in evidence-based treatments (EBT) alone does not result in therapist behavior change when providing psychotherapy. Rather, ongoing support and feedback is important (Sholomskas et al., 2005; Beidas & Kendall, 2010).

Accountability in Supervision

In order to maximize therapist application of evidence during treatment to improve service quality and effectiveness, it is critical to identify specific supervision strategies that positively influence therapist behavior during treatment. Previous studies have begun to delineate such strategies (e.g., supervisor modeling of therapeutic techniques, Bearman et al., 2013; role

play, performance feedback, Bearman et al., 2017; therapy tape review paired with performance feedback, Martino et al., 2016; Weck et al., 2013). However, literature regarding how and why supervision achieves changes in therapist behavior remain nascent. Findings are mixed regarding the extent to which supervision impacts clinical outcomes, the true "acid test" of quality supervision (Ellis & Ladany, 1997). For example, Callahan et al., 2009 found that supervisor accounted for 16% of the variance in client outcomes, representing a medium effect size, whereas Rousmaniere et al. (2014) found that supervisor accounted for less than 1% of the variance. As a first step in capitalizing on supervision to improve client outcomes, it is important to determine how supervision influences therapist behavior during treatment by directly testing supervision strategies.

Supervision behaviors that promote a sense of accountability, defined as perceived expectation and responsibility to fulfill certain duties or obligations, may represent key strategies by which clinical supervision can be used to promote therapist use of evidence (Burke & Saks, 2009). Much of the work on the role of clinical supervision focuses on strategies that foster clinical competency with the assumption that therapists will perform a clinical behavior when their skills have been adequately developed to do so (Bearman et al., 2013, Martino et al., 2016; Weck et al., 2013). In contrast, accountability adds a focus on creating expectations regarding performance of a clinical behavior (Burke & Saks, 2009).

Within the organizational management literature research has emerged seeking to support "transfer-of-training," the extent to which trainees generalize knowledge and skills learned in training to their performance on the job (Baldwin & Ford, 1988). Although there remains a lack of consensus regarding mechanisms of transfer in the literature, researchers commonly use Baldwin and Ford's (1988) three factor framework when investigating training inputs that result

in transfer: training design (e.g., content and format of initial training), trainee characteristics (e.g., self-efficacy and attitudes), and work environment (e.g., training supports, resources, and opportunities to apply training). Focusing on the post training inputs (trainee characteristics and work environment), in their review, Burke and Saks (2009) identified accountability as a critical factor that was implicit in many transfer-of-training strategies but was rarely identified explicitly. They argued that accountability underlies commonly identified components of training transfer including pay-for performance, performance feedback, and peer and supervisory support.

Building on this work, Grossman and Burke-Smalley (2018) described that accountability is fostered by strategies that strengthen the links between prescriptions, events, and identities. Prescription is defined as a directive to perform an expected behavior, event is defined as the context in which the behavior can be performed, and identity is defined as the personal qualities and skills of the actor of the behavior. Per Grossman and Burke-Smalley's framework, when a person is clear about the behavior they are expected to engage in (*prescription*) and the context in which they are expected to implement the behavior (*event*) and have the skills necessary to do so (*identity*), accountability is strong and follow-through execution of the expected behavior is likely. Research in the field of clinical psychology points to strategies that align with the accountability links proposed by Grossman and Burke-Smalley (2018) and can be readily applied in the context of clinical supervision. The present study aims to consider supervision strategies consistent with this accountability model and to test their influence on therapist follow through use of a behavior identified during supervision in the subsequent treatment session (herein referred to as the "expected practice"). Descriptions of the accountability links, review of relevant literature from clinical psychology, and examples in the context of clinical supervision will be presented. See Figure 1 for a diagram of the accountability

framework that informed the current study. For the purpose of manuscript clarity, examples from the literature and candidate supervisory behaviors will be mapped onto the accountability link for which they most closely align with the acknowledgement that certain strategies may, in reality, work to strengthen multiple accountability links. However, Grossman and Burke-Smalley emphasize the importance of strengthening *all* accountability links making precise matching of strategy to accountability link unnecessary.

Prescription-Event Link

The prescription-event link is considered strong when the individual is clear about the expectation to perform a behavior (prescription) in a given context (event). Goal setting represents an example of a strategy that can strengthen the prescription-identity link. There is a robust literature on the use of goal setting to promote behavior change across behaviors as varied as weight loss to energy conservation. Effective goal setting involves clearly operationalizing the expected behavior, monitoring progress, and setting a deadline for goal attainment. There also appears to be added effectiveness when goal progress is shared publicly. In their recent meta-analysis of the impact of monitoring progress on goal attainment, Harkin et al. (2016) found a medium pooled effect size of sharing about progress monitoring publicly on goal attainment. This effect was significantly larger than the effect of monitoring progress privately. Similarly, in their review of the impact of goal setting on behavior change, Epton et al. (2017) found that sharing about goals publicly was associated with changes in behavior.

The process of effective goal setting maps onto the prescription-event link the in the accountability model. Operationalizing the behavior that will move a person toward their goal creates a clear prescription *(*expectations regarding the use of skills and behaviors) while monitoring progress and setting a deadline strengthens the link between this prescription and

event. Especially pertinent to the use of supervision to foster accountability are the findings that public goal creation and monitoring increase the likelihood of behavior change. Although a person can feasibly create effective goals privately, it appears that the presence of a supportive other (e.g., supervisor), maximizes the benefit of goal setting. In the context of clinical supervision, a strategy that strengthens the prescription-event link could look like a supervisor saying to the therapist, "Great! So, next time you see this client for treatment (event), the plan is that you will provide psychoeducation about her depression diagnosis (prescription)."

Identity-Prescription Link

The identity-prescription link is considered strong when the individual perceives the expected behavior (prescription) as personally relevant to their role in the organization and possible to execute based on their skills and characteristics (identity). The process of therapy homework assignment and review point to principles of accountability and readily mapped onto Grossman and Burke-Smalley's (2018) model. It is common for therapists to assign between-session practice (i.e., homework) to clients to promote transfer of in-session skill learning into the contexts in which these skills will be needed (Kazantzis & Lampropoulos, 2002). This is not dissimilar from the aim for therapists to transfer skills learned in training (e.g., evidence-based therapeutic practices) into the context in which these skills will be needed (psychotherapy sessions). In their review of therapist homework assignment behaviors in session, Scheel et al., (2004) identified written records of planned homework, clear rationale for homework, and review of assigned homework in the next session as associated with client homework completion. Among these strategies, review of assigned homework in the next session appears to be particularly important for behavior change. In their study, Weck et al. (2013) used observational coding to assess the extensiveness of therapist homework assignment and review

and the impact of these strategies on client homework completion in the context of treatment for recurrent depression. Results of this study indicated that extensiveness of review of the previous sessions' homework assignment predicted client completion of their next homework assignment. However, extensiveness of homework assignment did not impact client homework completion. Findings were similar in the context of homework assignment in cognitive behavioral therapy treatment (CBT) for depression (Bryant et al., 1999) and in CBT for individuals with traumatic brain injury (Zelencich et al., 2020). Extensiveness of therapist homework review was predictive of subsequent client homework completion, whereas the extensiveness of therapist assignment was not. Therapists create the expectation for fulfillment of certain actions by the client through homework assignment and reinforce this expectation through review at the next session thus strengthening the identity-prescription link. Homework review may also present the opportunity for the therapist and client to adjust homework to better fit the client's life circumstances and skill level, thus increasing the personal relevance of the homework and further strengthening the identity-prescription link.

Provision of psychotherapy involves complex decision-making processes both within and between sessions (Jacinto et al., 2014). Supervisors seeking to support therapists to use evidence in treatment might begin by aiding in identification of opportunities for application of evidence and expressing the intention to follow up in the next supervision session (analogous to homework assignment). Review in the next supervision session might then reinforce the expectation for use of evidence and aid the therapist in prioritizing this practice in the face of competing demands. In the context of clinical supervision, a strategy that strengthens the identity-prescription link could look like a supervisor saying to the therapist, "Next time we meet, we can discuss how providing psychoeducation went during your treatment session and if

the client seemed to have a better understanding." During the next supervision it might also look like, "How did it go when you described symptoms and treatment for depression during your last treatment session with this client?" or finally, "How did it go when you used this practice in the past with other clients on your caseload?"

Identity-Event Link

The identity-event link is considered strong when individuals perceive themselves as efficacious (identity) to apply the expected behavior in a specific context (event). Active learning strategies, specifically modeling and role play, have been studied in the context of clinical supervision and have been shown to support therapist self-efficacy and to change therapist behavior in session. For example, in their study of community mental health therapists receiving supervision from expert, doctoral-level consultants, Bearman et al. (2013) found that supervisor modeling of therapeutic techniques was predictive of therapists' use of the practice planned for in supervision during their next psychotherapy session. Discussion of the planned practice alone, was not predictive of agreement between planned practice and use in next session. Building on this work, in Bearman et al.'s (2017) study, novice therapists received training on cognitive restructuring and were then randomized to a supervision-as-usual condition focused on activities common in community supervision (e.g., case conceptualization, therapeutic alliance, and case management) or an enhanced supervision condition that additionally involved active learning components (modeling, role-play of skills, and performance feedback). Although both groups demonstrated some improvement in fidelity to the intervention, competence, and CBT expertise, only therapists who received enhanced supervision demonstrated continued improvement across three sessions. These active learning strategies directly map onto the identity-event link that is strengthened when an individual feels efficacious to apply new skills. In the context of clinical

supervision, a strategy that strengthens the identity-event link could look like a supervisor saying to the therapist, “Let’s review the steps of motivational enhancement together and talk about any questions you have (identity) before your next session with this client (event).”

Study Aims

Guided by Grossman and Burke-Smalley’s (2018) model of accountability, the present study aimed to:

(1) Identify the presence of supervision strategies hypothesized to strengthen links between prescription, identity, and event for therapists in the context of clinical supervision focused on promoting client engagement in community-based mental health services. Describe the type and frequency of these supervision strategies through observational coding of supervision sessions.

(2) Investigate the association between supervision strategies hypothesized to strengthen accountability and subsequent therapist use of the expected practice during their next therapy session.

Data for this project come from a multi-site randomized trial examining how evidence is used to inform decisions made in clinical supervision to improve youth and families’ engagement in school-based mental health services. The following candidate supervision behaviors that align with the three accountability links were investigated:

- *Selects* | *Practice* (prescription-event link): identification of a clinical practice supported by evidence for increasing client engagement in treatment to be used by the therapist during their next treatment session.
- *Plans to Review* | *Practice* (prescription-identity link): discussion of intention to review therapist’s use of the expected practice during the next supervision meeting.

- *Reviews* | *Practice* (prescription-identity link): review of therapist's use of the expected practice during next supervision session.
- *Prepares for Application* | *Practice* (event-identity link): preparation to use the expected practice during the therapist's next treatment session through review of relevant materials, modeling, or role play.

It is hypothesized that each candidate supervision strategy will be uniquely associated with increased odds that therapists will perform the expected practice (i.e., practice discussed during the previous supervision session) during their next treatment session. It is also hypothesized that as more supervision strategies are used (i.e., all four strategies versus one strategy), the odds that the therapist will use the expected practice during their next treatment session will increase.

Methods

Participants

Participants for the present study were recruited as part of the Reaching Families Study, a multi-site cluster randomized trial examining the impact of a coordinated knowledge system on therapist use of evidence and subsequent client engagement in school-based mental health services. Therapists and supervisors recruited for this study were employed by either the Los Angeles Unified School District School Mental Health and Wellness Center Program in urban Los Angeles, California or by the South Carolina Department of Mental Health in rural South Carolina. Given differences in training and resources received by participants in the experimental versus control condition of the Reaching Families Study, only supervisors and their therapists in the experimental condition were included in the present study. As part of the Reaching Families Study, supervisor and therapist dyads audio recorded care episodes consisting of a sequence of supervision and treatment sessions for each client (i.e., $s1 \rightarrow t1 \rightarrow s2 \rightarrow t2$). Therapists worked

with clients for approximately four sessions before audio recording began meaning t1 represented the first *recorded* session rather than the first session of care. Audio recordings were transcribed and de-identified. There were instances in which therapists saw their clients twice between supervision meetings (i.e., s1→ t1→ t2→ s2) or dyads met twice for supervision prior to the next treatment session (i.e., s1→ t1→ s2→ s3→t2) due to scheduling challenges between both supervisor and therapist, and therapist and client. In all cases, treatment sessions and their immediately preceding supervision session were included with the purpose of examining the impact of supervision strategies on therapist use of the expected practice in the following treatment session. For example, in the case of s1→ t1→ s2→ s3→t2, the following sessions were retained: s1, t1, s3, t2). This selection criteria resulted in a sample of 192 treatment sessions and 192 supervision sessions conducted by 48 therapists overseen by 16 supervisors. Supervisors oversaw the work of between one and five therapists (M = 3.00, SD = 1.37). Therapists saw between one and four clients (M = 2.27, SD = 1.05) for one to two treatment sessions (see Table 1 for demographics and characteristics of supervisors and therapists).

Behavioral Observational Coding Systems

Supervision transcripts were coded using the Action Cycle and Use of Evidence Behavioral Observational Coding System (ACE-BOCS). ACE-BOCS is a modular coding system comprised of clinical reasoning activities. Therapy transcripts were then coded using the Use of Research Evidence Content Application: Engagement (URECA-E) Coding System which is a content specific modular coding system designed to interface with the broader ACE-BOCS system. These coding systems were not specifically designed to align with Grossman and Burke-Smalley's (2018) accountability model. However, several of the ACE-BOCS activity codes map readily onto the model of accountability. In addition, the ability to identify the expected clinical

practice from the supervision session (prescription) and compare it to the practices the therapist used in their next treatment session provide a unique opportunity to test the impact of accountability strategies in supervision on therapist behavior.

ACE-BOCS

ACE-BOCS is an observational coding system used in the Reaching Families Study to assess supervisor and therapist strategies that pertain to use of evidence (Park et al., 2020). Strategies were coded at the excerpt level (i.e., section of text within one supervision transcript, binary rating of presence or absence of the strategy) and event level (i.e., one supervision transcript, cumulative extensiveness rating of strategy performed across the entire event from 0, "not observed" to 5 "highly extensive execution of the strategy"). A team of graduate student raters trained to reliability using a series of "gold standard" rated transcripts created through consensus by ACE-BOCS developers. Once reliable with "gold standard" rated transcripts, raters coded transcripts independently and met weekly to discuss any questions and prevent coder drift.

The present study examined the following event-level codes identified as related to accountability, *Selects | Practice, Prepares for Application | Practice, Plans to Review | Practice,* and *Reviews | Practice.* In the pilot study of the ACE-BOCS coding system, reliability estimates of extensiveness ratings were in the excellent range (ICC > .90) for *Selects | Practice* and *Prepares for Application | Practice. Plans to Review | Practice* and *Review | Practice* did not occur across the 30 pilot transcripts. See Figure 2 for a diagram illustrating ACE-BOCS codes as they map onto the accountability links. In the present study, extensiveness ratings for the codes at the event level were converted to a binary rating of absent (i.e., extensiveness rating of 0) or present (i.e., extensiveness rating of 1 through 5) and paired with an object code. Object codes captured the referenced therapeutic practice, specifically practices identified as efficacious

for addressing client treatment engagement. For example, in a given supervision session, the supervisor and therapist might *Select | Psychoeducation about the Presenting Problem* or*, Select | Motivational Enhancement.* In the present study, object codes from ACE-BOCS were compared to the therapeutic practices used by therapists during the treatment session following supervision as captured by the URECA-E Coding System (described below).

Selects| Practice. The code, *Selects | Practice,* reflects supervisor and therapist identification of an expected practice to be used in the next treatment session and can include an explicit statement reflecting selection of the practice, weighing the pros and cons of the chosen practice over other practices or providing rationale for why a specific practice was selected. *Selects | Practice* relates to the prescription-event link in the accountability framework as it captures the expectation of therapist use of a specific practice (prescription) during a clinical encounter (event).

Prepares for Application | Practice. The code, *Prepares for Application | Practice,* reflects supervisor and therapist review and/or rehearsal of the expected practice in preparation for therapist use in the next treatment session and can include behaviors such as reviewing a description of the practice, role playing use of the practice, and modeling. *Prepares for Application | Practice* relates to the event-identity link in the accountability framework as it captures supervision strategies to increase the therapist's sense of self-efficacy (identity) in anticipation of using a practice in the next treatment session (event).

Plans to Review | Practice. The code, *Plans to Review | Practice,* reflects the stated intention to evaluate the therapist's use of the expected practice in the next supervision meeting and can include discussion of how therapist's performance will be evaluated in the coming supervision session. *Plans to Review | Practice* relates to the prescription-identity link in the

accountability framework as it captures the expectation of use of a specific practice (prescription) by the therapist (identity).

Reviews | Practice. The code, *Reviews | Practice,* reflects supervisor review of the therapist's use of the expected practice during a previous treatment session and can include comparison of therapist's performance to expected values (e.g., reviewing steps of the practice and comparing with the therapist's behaviors). *Reviews | Practice* also relates to the *prescription-identity* link in the accountability framework as it further emphasizes the expectation of therapist's use of a specific practice. This code was examined in addition to *Plans to Review | Practice* given research on psychotherapy homework assignment that found that review of homework was predictive of future client homework completion (e.g., Weck et al., 2013; Zelenchich et al., 2020; Bryant et al., 1999).

URECA-E

URECA-E is an observational coding system used to assess therapist use of 11 evidence-based treatment engagement practices during a treatment session. Therapist behaviors were coded at the excerpt level (i.e., section of text within one treatment transcript; binary rating of presence or absence of steps that comprise engagement practices) and event level (i.e., one treatment transcript). At the event level, steps that comprise engagement practices and can occur across multiple engagement practices were given an extensiveness rating from 0 – "absent" to 5 – "step was delivered with great depth and detail." The combination and extensiveness of steps that occurred in each therapy session were used to inform a 0 to 5 rating for the extensiveness of each possible engagement practice. A team of graduate student raters trained to reliably code transcripts using a series of "gold standard" rated transcripts created through consensus by

URECA-E developers. Once reliable with "gold standard" rated transcripts, raters coded transcripts independently and met weekly to discuss any questions and prevent coder drift.

For the purpose of analyses in the current study, event level practice codes were converted to a binary rating of absent (i.e., extensiveness rating of 0) or present (i.e., extensiveness rating of 1 through 5). URECA-E is an extension of the ACE-BOCS coding system. As such, object codes captured in the ACE-BOCS system correspond with practices coded using the URECA-E system. See Figure 2 for a diagram illustrating the mapping between ACE-BOCS and URECA-E codes. See Table 2 for reliability estimates for URECA-E and ACE-BOCS codes.

Data Analysis

Descriptive statistics were calculated for accountability codes in the supervision session and practice codes in the subsequent treatment session. Binary, multi-level logistic regression models were used to investigate the association between supervisory strategies and therapist use of the expected practice during the subsequent psychotherapy sessions. Because supervisors and therapists could discuss multiple practices (e.g., *Prepares for Application | Motivational Enhancement* and *Prepares for Application | Psychoeducation about Services*) and therapists could use multiple practices during treatment, data were formatted such that each unit (i.e., row in a data frame) of data represented the supervision and therapist behaviors related to each of the 11 engagement practices. Data were formatted such that practices were nested within supervision/treatment sessions, nested within students, nested within therapists, and nested within supervisors, resulting in a five-level nesting structure (see the supplementary file for example of data formatting). Prior to adding predictors of interest into the models, empty models with random intercepts for the five levels were run to examine variance explained by each level.

Levels that explained negligible variance in the outcome were removed prior to adding predictors and covariates.

In the first model, predictors of interest were the presence/absence of the four supervision strategies: *Selects | Practice, Prepares for Application | Practice, Plans to Review | Practice, and Reviews | Practice.* This model enabled exploration of the unique contribution of each of the four strategies over and above the other strategies on therapist use of the expected practice in the next treatment session. In the second model, the predictor of interest was a sum of the presence of the four strategies (ranging from 0, no supervision strategies were used to 4, all supervision strategies were used). This model enabled the exploration of the association between increasing emphasis on the use of an expected practice during supervision on therapist use of the expected practice in the next treatment session. This statistical model aligned with Grossman and Burke-Smalley's (2018) assertion that accountability is strongest when all links between prescriptions, events, and identity are strong. In both models, the following covariates were included: state (South Carolina or California), number of days between supervision and the next treatment session, and treatment sequence (i.e., first or second treatment session for a given student across time). State was included as a covariate due to known differences in state policies related to resources for mental health services, training for therapists, and caseload size. Number of days between supervision and the subsequent treatment session was included as the plan made during a supervision session before the next treatment session may have a diminished effect as more time passes between sessions and therapist memory of the plan degrades or priorities shift. Finally, treatment sequence was included given the possibility that discussion of an expected practice across two supervision sessions may result in greater follow through during the second treatment session relative to the first.

Results

Aim 1

The first aim of the study was to identify and describe the presence of supervision strategies hypothesized to strengthen links between prescription, identity, and event for therapists in the context of clinical supervision. Regarding *Selects* | *Practice*, at least one practice was selected for use in the next treatment session across 72.92% of supervision sessions. *Prepares for Application* | *Practice* occurred in more than half of supervision sessions (70.83%) and *Plans to Review* | *Practice* occurred in a minority (18.23%) of supervision sessions. However, the strategy, *Reviews* | *Practice*, occurred in more than half of supervision sessions (57.81%) meaning that even in the first coded supervision session, review was occurring for past sessions.

Time between supervision and the subsequent treatment session ranged from 0 to 63 days with an average of 7.93 days (*SD* = 9.63) between supervision and treatment. At least one expected practice (i.e., a practice that was discussed in supervision) occurred in 44.27% of treatment sessions. Regardless of correspondence between supervision and treatment, an engagement practice occurred in most treatment sessions (96.88%). The most used practice was *Psychoeducation about Services* (53.65%) followed by *Addressing Barriers to Treatment* (43.23%). The least commonly used practices were *Understanding Identities and Values* (11.98%) and *Motivational Enhancement* (14.06%).

Aim 2

A binary, multi-level logistic regression model was used to investigate if the presence of supervision strategies (*Selects* | *Practice, Prepares for Application* | *Practice, Plans to Review* | *Practice, Reviews* | *Practice*) predicted therapist performance of the expected practice in the next therapy session. Upon examination of the empty models with random intercepts for each of the five levels (level 1: practice, level 2: supervision/treatment sessions, level 3: students, level 4:

therapists, level 5: supervisors), the student and supervisor levels were removed as they explained were less that .00001 of variance in the outcome resulting in a three-level model (practices nested within supervision/treatment sessions, nested within therapists).

Results indicated that the supervision strategies *Selects | Practice*, *Prepares for Application | Practice,* and *Reviews | Practice* were associated with significantly increased odds of the therapist performing the expected practice in the next therapy session. Because all supervision strategies were included in the same model, odds ratios represent the unique contribution of each supervision behavior over and above the other supervision behaviors and covariates. When the supervision strategy, *Selects | Practice,* occurred, the odds of the therapist performing the expected practice in the subsequent therapy session was 2.94 times higher than when the behavior did not occur ($b = 1.08$, $p < .001$) When the supervision strategy *Prepares for Application | Practice* occurred, the odds of the therapist performing the expected practice in the subsequent therapy session was 3.10 times higher than when the strategy did not occur ($b = 1.13$, $p < .001$). Finally, when the supervision strategy *Reviews | Practice* occurred, the odds of the therapist performing the expected practice in the subsequent therapy session was 2.56 times higher than when the strategy did not occur ($b = .94$, $p <.001$). *Plans to Review | Practice* and included covariates did not significantly alter the odds of the therapist performing the expected practice in the subsequent therapy session. See Table 3 for the full model.

A second binary, multi-level logistic regression model was used to investigate the association between the sum of the four supervision strategies and therapist use of the expected practice in the subsequent therapy session. Results indicated that for every additional accountability strategy used, the odds of therapists performing the expected practice in the next therapy session increased by 2.80 ($b = 1.03$, $p <.001$). In other words, an increasing focus on a

specific practice as reflected by use of additional supervision strategies increased the likelihood the therapist would use that practice. Included covariates did not significantly increase the odds of the therapist performing the expected practice in the subsequent therapy session. See Table 4 for the full model.

Discussion

Guided by Grossman and Burke-Smalley's (2018) model of accountability from the organizational management literature, the aims of the present study were to identify candidate supervision strategies and test their influence on therapist use of evidence. Descriptive results indicated that the four candidate supervision strategies (*Selects* | *Practice, Prepares for Application* | *Practice, Plans to Review* | *Practice,* and *Reviews* | *Practice)* thought to strengthen the links between prescription, event, and identity did occur. However, the frequency of these behaviors across all supervision sessions were variable such that *Selects* | *Practice* occurred across most sessions (72.92%) whereas *Plans to Review* | *Practice* occurred in less than a quarter of sessions (18.23%).

Overall, results of analyses exploring the association between supervision strategies and therapist use of the expected practice in their next treatment session supported the accountability model. *Selects* | *Practice* (prescription-identity link), *Prepares for Application* | *Practice* (event-identity link), and *Reviews* | *Practice* (identity-prescription link) were uniquely associated with increased odds of therapist use of the expected practice in their next treatment session. However, *Plans to Review* | *Practice* was not associated with increased odds. In addition, the odds of therapist use of the expected practice increased as the number of supervision strategies used increased. There is a general sentiment in the research literature and field of clinical psychology at large that supervision is important for increasing service delivery quality, though research

directly testing this assumption is nascent. These findings are especially notable as they: 1) provide evidence that supervision can be used to influence clinical practice and 2) describe *how* supervision influences clinical practice through identification of specific supervision strategies. Findings from the present study point to three such behaviors supervisors can engage in that, independently and in combination, influence therapist behaviors in subsequent treatment sessions. An expanded discussion of each accountability link and associated supervision strategy follows.

Prescription-Event Link

The prescription-event link is considered strong when individuals are clear about the target context (i.e., treatment session) in which they are expected to use a specific practice (prescription). Results of the current study showed that when a practice was explicitly *selected* during supervision, the odds of therapist using the expected practice in the subsequent session were 2.94 times higher than when a practice was not selected. In the current study, it was not uncommon for supervisors and therapists to consider multiple practices in each meeting (M = 2.81, SD = 2.17). However, it was important for supervisors and therapists to narrow the focus and explicitly *select* the expected practice for the next session thus strengthening the link between the prescription and the event. And indeed, when it came to *selecting* a practice, supervision dyads did tend to narrow to one practice (M = 1.22, SD = 0.91). It may also be that setting a goal publicly (e.g., goal is discussed and settled upon by the supervisor and therapist) results in increased accountability relative to private goal setting (Harkin et al., 2016; Epton et al., 2017). However, future research is needed to determine if a coordinated knowledge system to plan for treatment used alone results in the same follow through as treatment plans made with the support of a clinical supervisor.

Identity-Prescription Link

The identity-prescription link is considered strong when individuals perceive the expected practice (prescription) as relevant to their role in the organization (identity). The process of setting an intention to review the therapist's use of an expected practice in the next supervision session (i.e., *Plans to Review | Practice*) and following through with the plan in the next supervision session (i.e., *Reviews | Practice)* were hypothesized to result in greater use of the expected practice in the next treatment session. *Reviews | Practice* did result in increased odds of therapist using the expected practice in the next treatment session. Contrary to the hypothesis, *Plans to Review | Practice* did not result in increased odds of therapist using the expected practice. Interestingly, this parallels findings in the therapy literature that thoroughness of therapist homework assignment does *not* result in greater client homework completion, however thoroughness of therapist homework review does result in greater client homework completion (Bryant et al., 1999, Weck et al., 2013, Zelencich et al., 2013). It may be that when supervisors and their therapists review previous use of a practice, they are able to troubleshoot the practice resulting in greater confidence of the therapist to implement the practice in their next session. Previous studies have emphasized the importance of corrective feedback for fostering clinical skill development (e.g., Martino et al., 2016, Webster-Stratton et al., 2014). Although not as intensive as corrective feedback provided through tape review as in Martino et al.'s (2016) study, discussion of past use of a practice may offer opportunities for feedback from the supervisor. It might also be, in reviewing therapist's use of the practice, the supervisor is able to effectively calibrate the preparation (e.g., modeling, role play, review of checklist) needed to ready the therapist for the next session.

Event-identity link

The event-identity link is considered strong when individuals perceive themselves as efficacious (identity) in implementing the expected practice in a specific context (event). Results showed that, when preparation for application of the practice occurred in supervision (i.e., *Prepare for Application | Practice),* the odds of therapist using the expected practice in their next therapy session were 3.10 times higher than when a practice was not prepared. In the present study, preparation could include several types of behaviors from discussion of the practice or review of a checklist describing the practice, to active modeling and role play. Previous studies have demonstrated the effectiveness of role plays for supporting therapist skill development (e.g., Bearman et al., 2013, Bearman et al., 2017, Martino et al., 2016). In contrast, in the present study, preparation broadly resulted in therapist follow through. It may be that supervision dyads are calibrating the type of preparation to the therapist's skill set and familiarity with the practice. For example, if a therapist has experience with a practice, a quick discussion of how to apply the practice when working with a specific client may be sufficient. In contrast, if a therapist has never used a practice, supervisor modeling of the practice may be warranted to build skill. An alternative explanation is that, in some cases, preparation serves to strengthen other aspects of accountability without increasing therapist self-efficacy to perform a skill (i.e., event-identity link). For example, discussion of how to apply a practice when working with a specific client may serve to clearly operationalize the prescription and anchor it to the event (i.e., prescription-event link).

Limitations

This study has several limitations. First, it is important to acknowledge following a plan identified in supervision is likely ineffective when not matched to the presenting problem of the client. There is both a need for congruence between presenting problem with plan made in

supervision *and* between plan made in supervision with therapist behavior in session. The present study focused in on the latter part with acknowledgement that upstream decisions are equally, if not more, important. Second, the coder reliability estimates were low for several of the codes (see Table 1). However, the full URECA and ACE-BOCS coding schemes included both event level codes (used in the current study) and excerpt level codes (i.e., line by line coding of transcribed text of the supervision and treatment session). Since the execution of this study, emerging analyses of event level codes derived from the excerpt level coding have shown that the latter approach is associated with greater reliability. These codes could, therefore, be used in analyses to increase confidence in present study results. In addition, the present study does not specify who initiated the accountability strategy during the supervision session. On the one hand, supervisor initiation may capitalize on the power dynamic in the supervisory dyad resulting in greater accountability to the authority figure. In contrast, therapist initiation may represent a more collaborative approach resulting in greater personal accountability and autonomy. Future studies exploring the moderating effect of the actor (supervisor or therapist) between supervision strategy and therapist use of the expected practice in the next treatment session are warranted.

The code, *Reviews | Practice,* was found to increase the odds of therapist use of that expected practice in the *next* treatment session. Given the design of this study, it is unclear whether therapist behavior in the previous session was driving therapist behavior in the next treatment session rather than the intervening supervision strategy itself. For example, if a therapist began use of the practice *Understanding Beliefs, Identities, and Values*, during t1, the supervisor may be prompted to check in to see how that went, thus receiving the code *Reviews | Understanding, Beliefs, Identities, and Values,* Logically, it is possible that the therapist might

finish that practice in t2 because they had already started it, regardless of supervision strategy used. However, this finding does mirror results from research related to therapist homework assignment and review (e.g., Bryant et al., 1999, Weck et al., 2013, Zelencich et al., 2013). To increase confidence in this finding as it pertains to supervision, future studies could consider creating an overall score for a supervisor's global propensity to review the plan made in the previous supervision session and use this score to explore the connection with therapist use of the expected practice.

Finally, the original coding schemes included extensiveness ratings of therapist and supervisor behaviors on a scale of 0 (behavior not observed) to 5 (highly extensive execution of the behavior). The structure of the data (i.e., each row of data represented behaviors performed in relation to a practice nested within a supervision to treatment sequence) resulted in a highly disproportionate number of zeros for the predictor variable as supervision strategies only occurred in relation to one or two practices. For this reason, models using extensiveness ratings failed to converge and instead this study used binary ratings of the presence/absence of supervision and therapist behaviors.

Future Directions

Findings from the current study point to several future directions. First, there was variable frequency of use of the 11 engagement practices during therapy sessions. *Psychoeducation about Services* was the most common (53.65% of sessions) and *Understanding Beliefs, Identities, and Values* (11.98% of sessions) was the least common. Results from previous research suggest that certain therapy practices tend to be used more frequently than others largely due to therapist-level factors (e.g., familiarity with intervention) rather than the relevance of the practice to the client (Starace, 2012; Kelley et al., 2010). For example, in their study of therapy

session content in usual care settings, Kelley et al (2010) found that 34% of the variance in practice selection was attributable to the clinician level compared with 11% at the client level. Future studies could explore this possibility by examining the agreement between the client's engagement difficulty and therapist intervention used to determine if there were missed opportunities for the less frequently used practices or if those practices were, in fact, less relevant to this sample of clients.

Relatedly, the models used in the present study did not account for the dependance among therapeutic practices. For example, *Prepares for Application | Psychoeducation about Services* was treated the same as *Prepares for Application | Motivational Enhancement*. It is possible that accountability strategies are more or less important depending on the characteristics of the expected therapeutic practice. For this reason, future studies might explore the interaction between accountability strategies and target therapeutic practices to determine if a greater degree of accountability is necessary for the execution of some practices but not others.

Given that some practices tended to be used more frequently than others, it would also be interesting to determine the extent to which supervision dyads adapt the intensity of their preparations during supervision. Although a therapist may benefit from intensive role play of a practice, they are less familiar with, it may be more efficient to quickly review a checklist describing the practice when they are more familiar. A future study might examine therapists' most used practices across all sessions and compare these to the strategies used in supervision to *prepare for application* to discern any attempts of supervisors to calibrate strategies to therapists' existing skills and knowledge. On a related note, a previous study (Boyd et al., 2021) found that it is not uncommon for therapists' skills and knowledge to differ from their supervisors. In instances where the therapists are familiar with more or different practices from their

supervisors, supervisors may perceive the relationship between themselves and their supervisees more negatively. For this reason, a future study might compare therapist and supervisor familiarity with each of the 11 practices and determine the extent of supervisor focus on specific practices relative to their comfort and familiarity with that practice.

Conclusions

The present study provides support for an organizational management model of accountability in which strengthening the links between prescriptions, identity, and event results in supervisee use of the expected therapeutic practice. Results provide evidence that supervision influences clinical practice and describe *how* supervision influences clinical practice through identification of specific supervision strategies*, Selects | Practice, Prepares for Application | Practice*, and *Reviews | Practice*.

Summative Discussion of Dissertation Findings

Clinical supervisors operate in strategic organizational positions in which they interface directly with upper-level management and therapists on the front line of care delivery. As such, there are opportunities for supervisors to shape therapists' organizational experience as well as provision of clinical services. This dissertation provides evidence of both. Results of the first study provide evidence that supervisors contribute to the perceptions of organizational climate of the therapists they oversee. Specifically, perceptions of interpersonal relationships with other employees (i.e., cohesion) and independence when performing work tasks (i.e., autonomy) were shared among therapists from the same supervision team and distinct from therapists on different supervision teams. Further, stronger supervisor communication was associated with greater therapist perceptions of cohesion and autonomy. Results of the second study demonstrate that supervision can directly influence therapist care provision through three strategies. Explicit

selection of a practice, *preparation* through activities such as reviewing manuals, conducting role plays, and modeling, and *review* of past use of the practice result in a nine-fold increase in the odds of therapists using an identified therapeutic practice during their next treatment session.

Together, findings from these studies point to the potential for supervisors to have a positive impact on the places they work and people they oversee. However, sufficient resources and support are needed to enable supervisors to maximize their role. First, at a minimum, supervisors should be included in trainings in evidence-based protocols and novel procedures that they are expected to supervise. Previous findings suggest that supervisors are sometimes trained in different or less evidence-based treatments compared with the therapists they oversee. The resulting training asymmetry has been shown to negatively impact relationships between supervisors and therapists and likely also hinders the supervisor's ability to support therapist provision of high-quality clinical interventions (Boyd et al., 2021). This is particularly troublesome given the consistent finding that ongoing support is necessary for therapists to sustainably implement new learning in their clinical practice (Sholomskas et al., 2005). Taken a step further, organizations may consider strategies modeled after the agency supervisor approach for scaling up evidence-based mental health services in community settings through rigorous training and consultation for supervisors in evidence-based care delivery. In this model, supervisors go on to train and support the therapists they oversee in their own agency (Southam-Gerow et al., 2014; Westman et al., 2020). This is distinct from the more traditional approach to implementing evidence-based care in which therapists receive didactic training in a discrete manualized treatment and are expected to implement it in their practice. The agency supervisor approach has been shown to be as effective as training and ongoing consultation from expert trainers, is more cost effective than traditional approaches, and makes use of supervisors'

organizational knowledge. Second, organizations may consider directing initiatives to cultivate leadership skill at the supervisor level and hire or promote individuals into supervisor positions based on their capacity for leadership. Past research has emphasized the role of leadership in improving organizational climate, strengthening capacity of therapists to form relationships with clients, and fostering positive attitudes towards evidence-based treatments (Aarons et al., 2011; Green et al., 2014; Brimhall et. al., 2016). Because supervisors provide leadership through oversight of therapists while also having direct face time, it may be particularly strategic to invest in leadership development at this level. Finally, sufficient time is needed for supervisors to perform their role and develop needed skill described above. Supervisors often carry additional responsibilities outside of direct oversite of therapist care delivery. For example, some supervisors in the Reaching Families study maintained their own clinical caseloads and held significant administrative responsibilities apart from supervision. Although having supervisors take on these tasks may be of short-term benefit for the organization, protected time to meet regularly with supervisees and to be responsive to emergent needs is likely important for maximizing the supervisory impact. The supervision role offers promising opportunities for improving organizational and individual capacity to provide high quality mental health treatment. To capitalize on this opportunity, intentional resources should be devoted to ensure providers have the needed skill and sufficient time to perform their role.

Figures

Figure 1. ACE-BOCS codes mapped onto accountability links

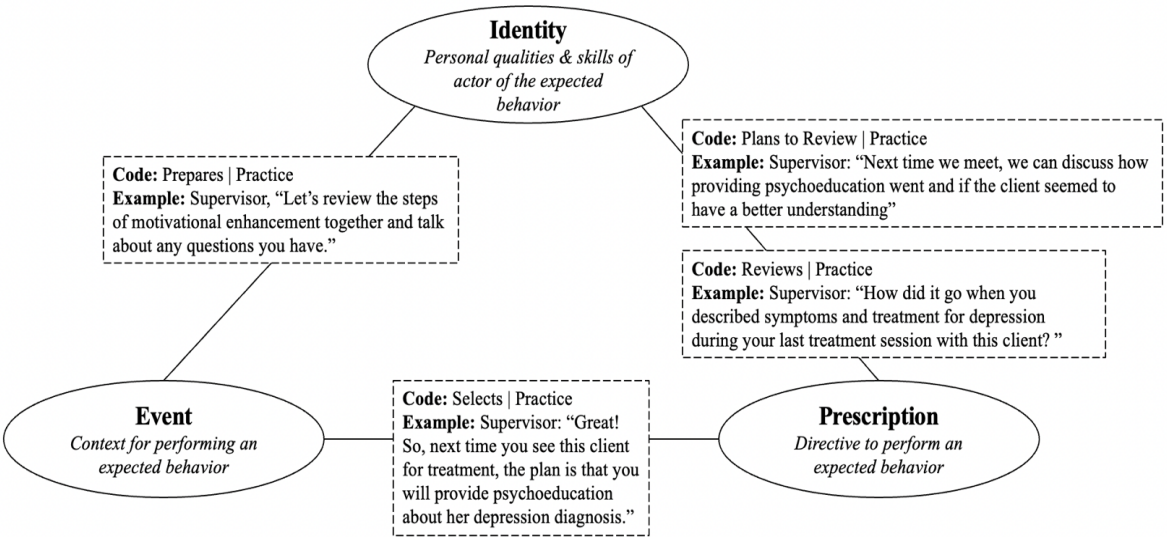
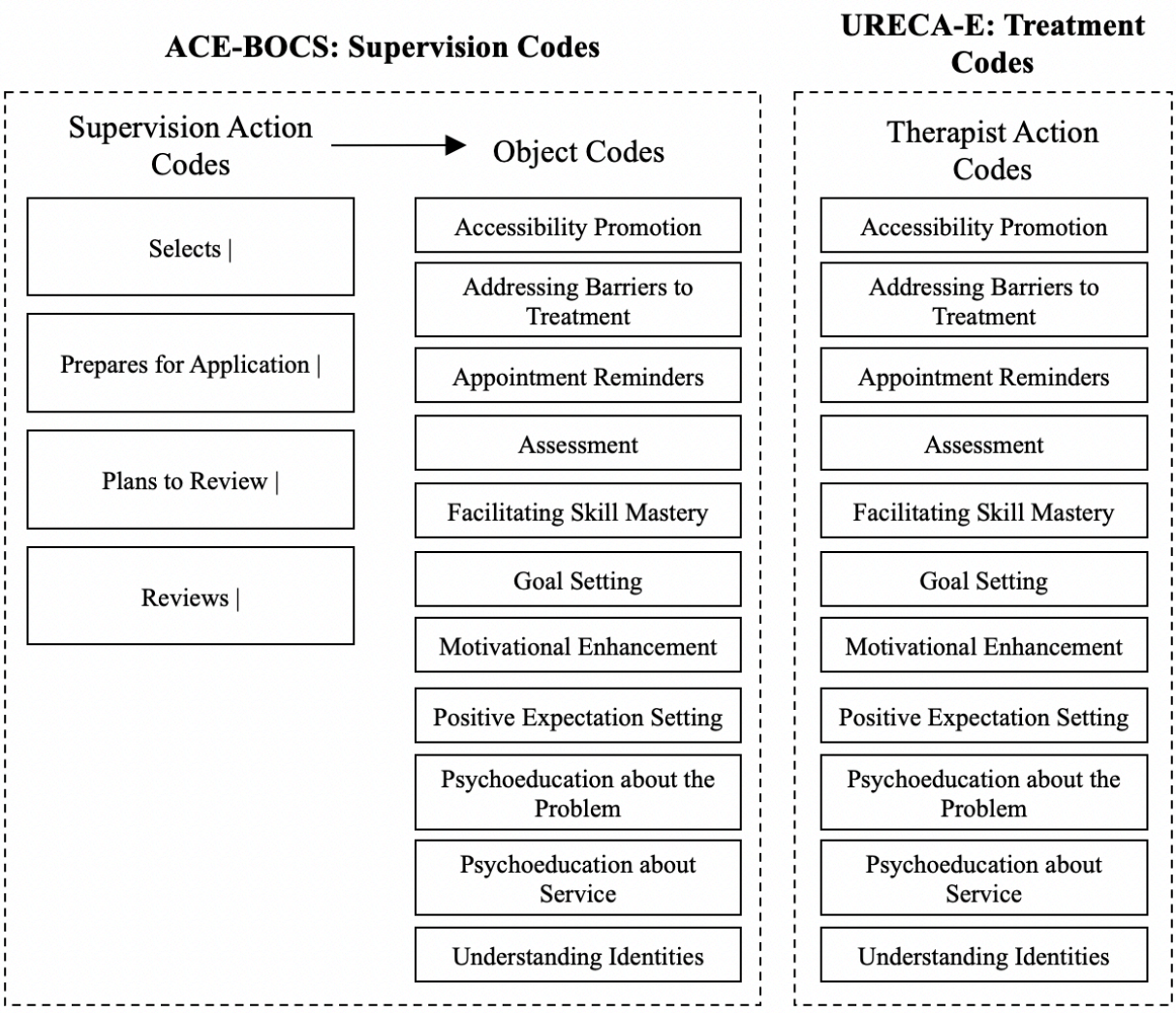


Figure 2. ACE-BOCS to URECA-E mapping



Tables

Table 10.

Demographics of Supervisor and Therapists

	Supervisor (n = 16)		Therapists (n = 48)	
	n	%	*n*	%
Race/Ethnicity				
Asian	0	0.00	2	4.17
Black/African American	8	50.00	20	41.67
White/Caucasian	4	25.00	4	8.33
Hispanic/Latinx	3	18.75	22	45.83
Multiracial	1	6.25	0	0.00
Gender				0.00
Woman	15	93.75	44	91.67
Man	1	6.25	4	8.33
State				0.00
California	9	56.25	28	58.33
South Carolina	7	43.75	20	41.67
Academic Degree				0.00
Master's-level	12	75.00	48	100.00
PhD	1	6.25	0	0.00
	M	*SD*	*M*	*SD*
Years of clinical experience	15.35	7.80	6.01	5.06

Note. Missing academic degree type and years of clinical experience from 3 supervisors

Table 11.

Agreement and Reliability Estimates for Supervision and Treatment Actions

	% Events	K	% Agreement
Supervision Codes			
Selects | Practice	58.54	0.39	71.95
Prepares for Application | Practice	53.66	0.66	82.93
Plans to Review | Practice	12.20	0.51	87.80
Reviews | Practice	74.39	0.83	93.90
Treatment Codes			
Accessibility Promotion	76.25	0.26	78.75
Addressing Barriers to Treatment	31.25	0.37	65.00
Appointment Reminders	16.25	0.51	85.00
Assessment	46.25	0.48	53.75
Facilitating Skill Mastery	43.75	0.75	65.00
Goal Setting	16.25	0.54	83.75
Motivational Enhancement	8.75	0.78	91.25
Positive Expectation Setting	11.25	0.28	76.25
Psychoeducation about Services	42.50	0.48	56.25
Psychoeducation about Problem	40.00	0.46	58.75
Understanding Identities	7.50	0.36	92.50

Note. Supervision $n = 82$, treatment $n = 80$; reliability estimates were calculated using supervision and treatment sessions from the experimental and control condition of the Reaching Families Study

Table 12.

Effect of Presence of Supervision Behaviors on Therapist use of Expected Practice

Factor	*b*	*SE*	Odds Ratio	*p*
Supervision behaviors				
Selects | Practice	1.08	0.21	2.94	<.01
Prepares for Application | Practice	1.13	0.22	3.10	<.01
Plans to Review | Practice	0.60	0.44	1.82	0.17
Reviews | Practice	0.94	0.18	2.56	<.01
Covariates				
State (South Carolina)	-0.24	0.16	0.79	0.14
Days between supervision and treatment	-0.01	0.01	0.99	0.44
Treatment session count	-0.15	0.11	0.86	0.17

Variance Components	Variance	*SD*
Therapist (Intercept)	0.10	0.31
Session (Intercept)	0.14	0.38

Note. Therapists (n = 48), Sessions (n = 192); California was the reference group for State

Table 13.

Effect of Count of Supervisor behaviors on Therapist use of Expected Practice

Factor	*b*	*SE*	Odds Ratio	*p*
Count of supervision behaviors	1.03	0.08	2.8	<.01
Covariates				
State (South Carolina)	-0.23	0.16	0.79	0.15
Days between supervision and treatment	-0.01	0.01	-	0.44
Treatment session count	-0.16	0.11	0.85	0.14
Variance Components	Variance	*SD*		
Therapist (Intercept)	0.10	0.31		
Session (Intercept)	0.15	0.38		

Note. Therapists (n = 48), Sessions (n = 192); California was the reference group for State

References

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, *41*(1), 63-105. https://doi.org/10.1111/j.1744-6570.1988.tb00632.x

Bailin, A., Bearman, S. K., & Sale, R. (2018). Clinical supervision of mental health professionals serving youth: Format and microskills. *Administration and Policy in Mental Health and Mental Health Services Research*, *45*(5), 800-812. https://doi.org/10.1007/s10488-018-0865-y

Bearman, S. K., Weisz, J. R., Chorpita, B. F., Hoagwood, K., Ward, A., Ugueto, A. M., ... & Research Network on Youth Mental Health. (2013). More practice, less preach? The role of supervision processes and therapist characteristics in EBP implementation. *Administration and Policy in Mental Health and Mental Health Services Research*, *40*(6), 518-529. https://doi.org/10.1007/s10488-013-0485-5

Bearman, S. K., Schneiderman, R. L., & Zoloth, E. (2017). Building an evidence base for effective supervision practices: An analogue experiment of supervision to increase EBT fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*(2), 293-307. https://doi.org/10.1007/s10488-016-0723-8

Beidas, R. S., & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, *17*(1), 1-30. https://doi.org/10.1111/j.1468-2850.2009.01187.x

Boyd, M. R., Park, A. L., Becker, K. D., & Chorpita, B. F. (2021). The relation between training asymmetry and supervisory working alliance: Implications for the role of supervisors in

implementation. *The Clinical Supervisor*, *40*(1), 49-67. https://doi.org/10.1080/07325223.2020.1871460

Bryant, M. J., Simons, A. D., & Thase, M. E. (1999). Therapist skill and patient variables in homework compliance: Controlling an uncontrolled variable in cognitive therapy outcome research. *Cognitive Therapy and Research*, *23*(4), 381-399. https://doi.org/10.1023/A:1018703901116

Burke, L. A., & Saks, A. M. (2009). Accountability in training transfer: Adapting Schlenker's model of responsibility to a persistent but solvable problem. *Human Resource Development Review*, *8*(3), 382-402. https://doi.org/10.1177/1534484309336732

Callahan, J. L., Almstrom, C. M., Swift, J. K., Borja, S. E., & Heath, C. J. (2009). Exploring the contribution of supervisors to intervention outcomes. *Training and Education in Professional Psychology*, *3*(2), 72. https://doi.org/10.1037/a0014294

Dorsey, S., Pullmann, M. D., Deblinger, E., Berliner, L., Kerns, S. E., Thompson, K., ... & Garland, A. F. (2013). Improving practice in community-based settings: a randomized trial of supervision–study protocol. *Implementation Science*, *8*(1), 1-11. https://doi.org/10.1186/1748-5908-8-89

Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision: An integrative review. In C. E. Watkins, Jr. (Ed.), *Handbook of Psychotherapy Supervision* (pp. 447–507). John Wiley & Sons, Inc.

Epton, T., Currie, S., & Armitage, C. J. (2017). Unique effects of setting goals on behavior change: Systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, *85*(12), 1182. https://doi.org/10.1037/ccp0000260

Gray, M., Sharland, E., Heinsch, M., & Schubert, L. (2015). Connecting research to action: Perspectives on research utilisation. *The British Journal of Social Work*, *45*(7), 1952-1967. https://doi.org/10.1093/bjsw/bcu089

Grossman, R., & Burke-Smalley, L. A. (2018). Context-dependent accountability strategies to improve the transfer of training: A proposed theoretical model and research propositions. *Human Resource Management Review*, *28*(2), 234-247. https://doi.org/10.1016/j.hrmr.2017.08.001

Harkin, B., Webb, T. L., Chang, B. P. I., Prestwich, A., Conner, M., Kellar, I., Benn, Y., & Sheeran, P. (2016). Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychological Bulletin, 142*(2), 198–229. https://doi.org/10.1037/bul0000025

Jacinto, S. B., Lewis, C. C., Braga, J. N., & Scott, K. (2018). A conceptual model for generating and validating in-session clinical judgments. *Psychotherapy Research*, *28*(1), 91-105. https://doi.org/10.1080/10503307.2016.1169329

Kazantzis, N., & Lampropoulos, G. K. (2002). Reflecting on homework in psychotherapy: What can we conclude from research and experience? *Journal of Clinical Psychology*, *58*(5), 577-585. https://doi.org/10.1002/jclp.10034

Kelley, S. D., de Andrade, A. R. V., Sheffer, E., & Bickman, L. (2010). Exploring the black box: Measuring youth treatment process and progress in usual care. *Administration and Policy in Mental Health and Mental Health Services Research*, *37*(3), 287-300. https://doi.org/10.1037/12075-011

Martino, S., Jr.Paris, M., Añez, L., Nich, C., Canning-Ball, M., Hunkele, K., & ... Paris, M. J. (2016). The effectiveness and cost of clinical supervision for motivational interviewing:

A randomized controlled trial. *Journal of Substance Abuse Treatment*, 68, 11–23. https://doi.org/10.1016/j.jsat.2016.04.005

Park, A. L., Becker, K. D., Boustani, M. M., & Chorpita, B. F. (2020). Decision-making in mental health care: Measuring provider and supervisor use of evidence. *Administration and Policy in Mental Health and Mental Health Services Research*, 1-13. https://doi.org/10.1007/s10488-019-00989-0

Powell, B. J., Waltz, T. J., Chinman, M. J., Damschroder, L. J., Smith, J. L., Matthieu, M. M., ... & Kirchner, J. E. (2015). A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implementation Science*, *10*(1), 1-14. https://doi.org/10.1186/s13012-015-0209-1

Rousmaniere, T. G., Swift, J. K., Babins-Wagner, R., Whipple, J. L., & 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

Schoenwald, S. K., Chapman, J. E., Kelleher, K., Hoagwood, K. E., Landsverk, J., Stevens, J., Glisson, C., Rolls-Reutz, J. & Research Network on Youth Mental Health. (2008). A survey of the infrastructure for children's mental health services: Implications for the implementation of empirically supported treatments (ESTs). *Administration and Policy in Mental Health and Mental Health Services Research*, *35*(1-2), 84-97. https://doi.org/10.1007/s10488-007-0147-6

Sholomskas, D. E., Syracuse-Siewert, G., Rounsaville, B. J., Ball, S. A., Nuro, K. F., & Carroll, K. M. (2005). We don't train in vain: a dissemination trial of three strategies of training clinicians in cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology*, *73*(1), 106. https://doi.org/10.1037/0022-006X.73.1.106

Scheel, M. J., Hanson, W. E., & Razzhavaikina, T. I. (2004). The process of recommending homework in psychotherapy: A review of therapist delivery methods, client acceptability, and factors that affect compliance. *Psychotherapy: Theory, Research, Practice, Training*, *41*(1), 38. https://doi.org/10.1037/0033-3204.41.1.38

Southam-Gerow, M. A., Daleiden, E. L., Chorpita, B. F., Bae, C., Mitchell, C., Faye, M., & Alba, M. (2014). MAPping Los Angeles County: Taking an evidence-informed model of mental health care to scale. *Journal of Clinical Child & Adolescent Psychology*, *43*(2), 190-200. https://doi.org/10.1080/15374416.2013.833098

Starace, N. (2012). *Trauma Translated into Therapeutic Practice: Posttraumatic Stress Disorder and Correspondence with Treatment Target Selection in a Usual Care Setting*. [Unpublished doctoral dissertation]. University of California, Los Angeles.

Webster-Stratton, C. H., Reid, M. J., & Marsenich, L. (2014). Improving therapist fidelity during implementation of evidence-based practices: Incredible years program. *Psychiatric Services*, *65*(6), 789-795. https://doi.org/10.1176/appi.ps.201200177

Weck, F., Richtberg, S., Esch, S., Höfling, V., & Stangier, U. (2013). The relationship between therapist competence and homework compliance in maintenance cognitive therapy for recurrent depression: Secondary analysis of a randomized trial. *Behavior Therapy*, *44*(1), 162-172. https://doi.org/10.1016/j.beth.2012.09.004

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: a meta-analysis of direct comparisons. *American Psychologist*, *61*(7), 671. https://doi.org/10.1037/0003-066X.61.7.671

Weisz, J. R., Kuppens, S., Eckshtain, D., Ugueto, A. M., Hawley, K. M., & Jensen-Doss, A. (2013). Performance of evidence-based youth psychotherapies compared with usual

clinical care: A multilevel meta-analysis. *JAMA Psychiatry*, *70*(7), 750-761. https://doi.org/10.1001/jamapsychiatry.2013.1176

Westman, J. G., Daleiden, E. L., & Chorpita, B. F. (2020). The agency supervisor model: developing supervisors who facilitate therapist transfer of training in community behavioral health service organizations. *The Clinical Supervisor*, *39*(1), 45-65. https://doi.org/10.1080/07325223.2019.1695159

Zelencich, L. M., Wong, D., Kazantzis, N., McKenzie, D. P., Downing, M., & Ponsford, J. L. (2020). Predictors of anxiety and depression symptom improvement in CBT adapted for traumatic brain injury: Pre/post-Injury and therapy process factors. *Journal of the International Neuropsychological Society: JINS*, *26*(1), 97-107. https://doi.org/10.1017/S1355617719000791