

UCLA

UCLA Electronic Theses and Dissertations

Title

Therapist Missteps in Youth Mental Health Services

Permalink

<https://escholarship.org/uc/item/2xz8x8ph>

ISBN

9798184125442

Author

Park, Hyun Seon

Publication Date

2026-07-21

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Therapist Missteps in Youth Mental Health Services

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy in Psychology

by

Hyun Seon Park

2026

© Copyright by

Hyun Seon Park

2026

ABSTRACT OF THE DISSERTATION

Therapist Missteps in Youth Public Mental Health Services

by

Hyun Seon Park

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2026

Professor Bruce Frederick Chorpita, Chair

Treatment engagement challenges are among the most persistent barriers to effective youth mental health treatment, with a substantial proportion of youth and families terminating services before achieving expected clinical gains (Becker & Chorpita, 2023). Despite growing recognition of the multidimensional, dynamic, and transactional nature of treatment engagement, relatively little is known about the within-session therapist behaviors that may occur in the context of and inadvertently contribute to engagement challenges. This dissertation examined therapist “missteps,” or inadvertent, less optimal, or noncollaborative therapist behaviors, across two studies using observational data from the Reaching Families Study, a multisite cluster-randomized trial focused on improving use of evidence to target low treatment engagement in publicly funded youth mental health settings (Chorpita et al., 2025).

Study 1 sought to first understand *how* therapists misstep by examining the occurrence of missteps across several key contextual factors, including session composition, the recipient of the misstep, and the surrounding in-session therapist behaviors. Therapist missteps were prevalent,

occurring in over half of recorded therapy sessions with an average of approximately 2.5 missteps per session. In family sessions, missteps were most frequently directed at the youth. Advice-giving emerged as the most prevalent type of therapist misstep to occur across all session types. Descriptively, approximately 38.1% of observed missteps occurred following another misstep occurrence (10.6%) or in the absence of an identifiable evidence-based technique (27.5%). These findings suggest that missteps may be a common occurrence when delivering care, particularly in the context of engagement challenges, and may be more vulnerable to occur during unstructured or uncertain moments in a session.

The first study also sought to examine *when* therapists may misstep by examining associations between misstep occurrences and factors related to therapist preparedness. Findings suggest that neither the timing of supervision relative to treatment nor the presence and quality of a treatment plan established for the session were credibly associated with total misstep frequency. These findings converge with later findings in Study 2, which indicated that the presence of a treatment plan did not appear to moderate the relationship between misstep frequency and average practice extensiveness, an aggregate indicator of the overall quality of treatment delivered in the session. Together, these findings suggest that identifying what to do in the next session alone may not be sufficient in supporting therapists from engaging in inadvertent missteps. However, provision of comprehensive, coordinated decision-making supports for the detection, identification, and targeted intervention of engagement challenges (i.e., the Coordinated Knowledge System, CKS) was associated with substantially lower expected misstep rates in Study 1. This robust effect remained in exploratory analyses even after accounting for the therapist's optimal practice quality in the session (i.e., the maximum practice extensiveness across all practices delivered in the session).

Study 2 extended these findings by examining what happens in the therapy session when missteps occur. Misstep frequency was credibly associated (i.e., in Bayesian estimation, displaying sufficient evidence that the effect is real and in a consistent direction given the posterior distribution of the parameter) with the therapist's overall treatment quality for the session (i.e., average practice extensiveness) in the originally proposed analyses, however, the credible interval marginally excluded zero and should be interpreted with caution. Exploratory analyses using the therapist's maximum practice extensiveness revealed a credible association with misstep occurrences, in which increases in the therapist's practice quality were associated with decreases in misstep occurrences in the session. Misstep occurrences were also credibly associated with a modest decrease in session-level engagement for both youth and caregivers, after accounting for baseline engagement challenges, maximum practice extensiveness, site, and study condition. Variance decomposition of youth and caregiver session-level engagement also revealed that although the therapist level accounted for significant variance (23.0%) in youth session engagement, caregiver session engagement was more organized at the client level (22.3%). Together, these findings reinforce the conceptualization of engagement as a dynamic, state-like phenomenon that is co-constructed within the session rather than trait-like features of clients, and position therapist behaviors as an important component of session-to-session engagement.

Overall, these studies found that therapist missteps occur routinely in youth publicly funded mental health settings and are associated with measurable changes in session-level engagement. Additionally, provision of coordinated clinical supports aimed at supporting therapists with navigating well-documented "blind curves" when delivering care in complex contexts, as well as extensive delivery of an evidence-based practice, were associated with

expected decreases in misstep occurrences. These findings continue to support the theoretical framing of therapist missteps as predictable features of treatment delivery and center a locus of intervention that shifts away from the therapist's intrinsic traits and qualities and towards scaffolding flexible, matched, and coordinated clinical decision-making supports to improve care for youth and families.

The dissertation of Hyun Seon Park is approved.

Kimberly D. Becker

Denise A. Chavira

Han Du

Bruce Frederick Chorpita, Committee Chair

University of California, Los Angeles

2026

Dedication

This dissertation would not have been possible without the guidance, mentorship, care, and wisdom of many individuals who have shaped my training as a researcher and clinician.

First, I would like to express my deepest gratitude to my advisor and dissertation chair, Dr. Bruce Chorpita. Thank you for taking a chance on me as your graduate student. You have helped me grow in countless ways over the past six years, and I will carry forward all I have learned under your guidance wherever I go.

I would also like to thank my committee member, Dr. Kim Becker, for her generosity, knowledge, and mentorship over the past six years, which have been instrumental to my growth. I also thank my other committee members Drs. Denise Chavira and Han Du, for their time, thoughtful feedback, and guidance, which strengthened this work in valuable ways.

To the clients, families, clinicians, supervisors, and community partners I have had the privilege of working with and learning from—you have taught me more than I can fully describe, and I am deeply grateful for the wisdom you have shared with me. I thank each and every one of you.

To my community of friends and loved ones—thank you for the strength, laughter, joy, and space you have created for me throughout this journey. The warmth you have gifted me is something I cherish deeply.

To Kumo, my cat, for your steadfast company from the first year of this work through its conclusion.

To my family—my parents, 박종태 and 박영선, and my brother, Sejun Park: I dedicate this to you. I am who I am today because of you. 나를 사랑해주고, 응원해주고, 내가 할수 있다고 믿어줘서 진심으로 감사합니다.

Table of Contents

GENERAL ABSTRACT OF THE DISSERTATION	iii
Acknowledgements	x
Chapter 1	1
Introduction	2
Therapist Missteps	3
Challenges with Effective Management of Engagement	5
Therapist Preparedness and “Run Time” Decisions	9
The Present Study	12
Methods	14
Study Context	14
Participants	15
Study Condition	16
Site	17
Measures	17
Data Analytic Plan	20
Results	24
Aim 1: Descriptive Characterization of Therapist Missteps	24
Aim 2: Effects of Therapist Preparedness on Frequency of Therapist Missteps	26
Discussion	28
How Therapists Misstep	29
When Therapists Misstep	31
Limitations	34
Chapter 2	37
Introduction	38
Missteps and the Quality of Treatment Delivery	39
Missteps and Associations with Engagement Outcomes	41

The Present Study	43
Methods.....	45
Study Context and Participants	45
Measures	46
Data Analytic Plan	49
Results.....	57
Aim 1: Effect of Therapist Missteps on Average Practice Quality	57
Exploratory Analyses: Effect of Maximum Practice Quality on Therapist Missteps.....	59
Aim 2: Effect of Therapist Missteps on Youth and Caregiver Session Engagement	59
Discussion	62
Treatment Planning	62
Session Engagement	63
Maximum Practice Extensiveness	65
Therapist Missteps	67
Study Condition	68
Limitations	69
Conclusion	71
Tables	73
Figures.....	96
Appendix A.....	101
Appendix B	103
Appendix C	111
References.....	113

Acknowledgements

Portions of this dissertation have been previously published:

Park, H. S., Becker, K. D., Chu, W., Reeder, K., Lechuga, C., Ahuna, J. K., ... Chorpita, B. F.

(2025). Identification of Potential Therapist “Missteps” in the Context of Engagement in Public Mental Health Services: Analysis of Digital Recordings and Key Informant Validation. *Psychotherapy Research*, 1–15.

<https://doi.org/10.1080/10503307.2025.2521792>

Park, H. S., Becker, K. D., & Chorpita, B. F. (2025). Understanding How and When Therapists Misstep: Navigating Engagement Challenges in Youth Mental Health Services. *Research on child and adolescent psychopathology*, 53(5), 669–685.

<https://doi.org/10.1007/s10802-025-01297-y>

HYUN SEON PARK, M.A., C.Phil.

Curriculum Vitae

EDUCATION

- University of California, Los Angeles (UCLA)** 2020-2026
Ph.D. Candidate in Clinical Psychology
Academic Advisor: Bruce F. Chorpita, Ph.D.
Dissertation: *Therapist Missteps in Youth Mental Health Services*
- University of California, Los Angeles (UCLA)** 2020-2022
Master of Arts in Clinical Psychology
Academic Advisor: Bruce F. Chorpita, PhD
Master's Thesis: *Contraindications in School-Based Mental Health Services*
- University of Illinois, Urbana-Champaign (UIUC)** 2014-2018
Bachelor of Sciences in Psychology with Clinical/Community Concentration; English

FELLOWSHIPS & AWARDS

- UCLA Psychology Dissertation Fellowship (\$10,000)** 2024
Eudaimonia Award 2024
Awarded by UCLA Semel Healthy Campus Initiative Center in recognition of contributions to creating a culture of caring, commitment to the good of society, and for exemplifying a life full of purpose and meaning.
- Graduate Summer Research Mentorship Program (\$6,000)** 2021
For research activities conducted collaboratively with faculty mentor (Bruce F. Chorpita, Ph.D.).
Role: PI. Awarded by UCLA Graduate Division
- Graduate Summer Research Mentorship Program (\$6,000)** 2020
For research activities conducted collaboratively with faculty mentor (Bruce F. Chorpita, Ph.D.).
Role: PI. Awarded by UCLA Graduate Division
- Illinois Program for Research in the Humanities Fellowship (\$5,000)** 2017
Awarded by UIUC Humanities Research Institute.

PEER-REVIEWED PUBLICATIONS

- Wu, E. G., Becker, K. D., Isenberg, B. M., **Park, H. S.**, Ahuna, J. K., Arkin, S., & Chorpita, B. F. (2026). Gaps in Detection: Comparing Provider- and Client-Reported Indicators of Engagement in Youth Telemental Health. *Journal of Clinical Child & Adolescent Psychology*, 1–14. <https://doi.org/10.1080/15374416.2026.2626905>
- Wu, E. G., Becker, K. D., Chu, W., Ahuna, J. K., Lakind, D., Boyd, M. R., Reeder, K., **Park, H. S.**, Knudsen, K. S., Guan, K., Arkin, S., Bradley, W. J., & Chorpita, B. F. (2026). Expanding Integrity Measurement in Mental Health Care: A Coding System of Targets, Practices, and Steps. *Journal of clinical child and adolescent psychology : the official journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, 1–19. Advance online publication. <https://doi.org/10.1080/15374416.2025.2587175>
- Park, H. S.**, Becker, K. D., Chu, W., Reeder, K., Lechuga, C., Ahuna, J. K., ... Chorpita, B. F. (2026). Identification of Potential Therapist “Missteps” in the Context of Engagement in Public Mental Health Services: Analysis of Digital Recordings and Key Informant Validation. *Psychotherapy Research*, 36(4), 674–688. <https://doi.org/10.1080/10503307.2025.2521792>
- Park, H.S.**, Becker, K. D., & Chorpita, B. F. (2025). Understanding How and When Therapists Misstep: Navigating Engagement Challenges in Youth Mental Health Services. *Research on child and adolescent psychopathology*, 53(5), 669–685. <https://doi.org/10.1007/s10802-025-01297-y>
- Isenberg, B.M., Becker, K.D., Wu, E., **Park, HS.**, Chu, W., Keenan-Miller, D., & Chorpita, B.F. (2023). Toward efficient, sustainable, and scalable methods of treatment characterization: An investigation

of coding clinical practice from chart notes. *Administration and Policy in Mental Health and Mental Health Services Research*. <https://doi.org/10.1007/s10488-023-01316-4>

- Moitra, E., **Park, H.S.**, Guthrie, K.M., Peters, G., Wittler, E., Gaudiano, B.A. (2022). Stakeholder perspectives on adjunctive mHealth services during transitions of care for patients with schizophrenia-spectrum disorders. *Journal of Mental Health*.
- Uebelacker, L.A., Wolff, J.C., Guo, J., Conte, K., Segur, R., Caviness, C.M., **Park, H.S.**, Peterson, S., Tremont, G., Rosen, R.K., Yen, S. (2021). Single-arm pilot trial of yoga for adolescents with depression. *Clinical Child Psychology and Psychiatry*.
- Moitra, E. **Park, H.S.**, Gaudiano, B.A. (2021). Development and initial testing of an mHealth transitions of care intervention for adults with schizophrenia-spectrum disorders immediately following a psychiatric hospitalization. *Psychiatric Quarterly*.
- Park, H.S.**, Moitra, E., Gaudiano, B.A. (2020). Lessons Learned When Using Mobile Ecological Momentary Assessment in Patients with Psychotic-Spectrum Disorders Following a Psychiatric Hospitalization. *SAGE Research Methods Cases*.
- Moitra, E., **Park, H.S.**, Ben-Zeev, D., & Gaudiano, B.A. (2020). Using ecological momentary assessment for patients with psychosis post-hospitalization: Opportunities for mobilizing measurement-based care. *Psychiatric Rehabilitation Journal*.

TALKS & PRESENTATIONS (* indicates mentee author)

- Chu, W., **Park, H.S.**, Arkin, S. (2023). Contextual and Transactional (and Surprising) Factors Related to Treatment Engagement. 57th Annual Association of Behavioral and Cognitive Therapies (ABCT) Convention, Seattle, WA.
- Reeder, K., **Park, H.S.**, Becker, K. D., & Chorpita, B. F. (2022). What drives therapist and supervisor evidence use? Exploring self-perceptions of use and attitudes toward evidence. 56th Annual ABCT Convention, Dissemination and Implementation Science Special Interest Group Exposition, New York, NY.
- Reeder, K., **Park, H.S.**, Becker, K. D., & Chorpita, B. F. (2022). Is training enough? Examining drivers of provider knowledge of evidence-based engagement practices. 6th Biennial SIRC Conference, San Diego, CA.
- Guo, A. *, **Park, H.S.**, Boyd, M., Becker, D., Chorpita, B. (2022, November). Therapist Burnout and Stress: Associations with Potentially Risky Behaviors in Therapy. 56th Annual ABCT Convention, New York, New York.
- Park, H.S.**, Reeder, K., Gugava, L. *, Becker, K., Chorpita, B. (2022, November). Measuring potentially risky behaviors in youth school-mental health services. 56th Annual ABCT Convention, New York, New York.
- Park, H.S.**, Reeder, K., Guo, A., Becker, K., Chorpita, B. (2021, November). Contraindications in therapy: Impact on Youth and Family Engagement in School-Based Mental Health Services. 55th Annual ABCT Convention, Virtual Convention
- Park, H.S.**, Peters, G., Wittler, E., Vining, J., Gaudiano, B.A., Moitra, E. (2020, November). Stakeholders' perspectives on using mobile technology to support patients with psychosis post-hospitalization. 54th Annual ABCT Convention, Virtual Convention.
- Park, H.S.**, Gaudiano, B.A., Moitra, E. (2019, March). Usability of electronic medication adherence monitoring (MEMS®) devices in patients with psychotic-spectrum disorders immediately following a psychiatric hospitalization. 22nd Annual Mind Brain Research Day, Brown University, Providence, Rhode Island.
- Park, H.S.**, Yeary, H., Halewicz, V.E. (2018, April). *Unwinding Time and Temporality*. Symposium presented at Illinois Program for Research in the Humanities Undergraduate Symposium, Champaign, IL.

Chapter 1

Understanding How and When Therapists Misstep While Navigating Engagement Challenges in Youth Mental Health Services

Introduction

Treatment engagement—or the ability to fully participate in and benefit from services—is a critical component of effective youth mental health care (Becker & Chorpita, 2023). Between 50% and 70% of youth and families terminate services earlier than recommended and prior to achieving expected gains from treatment (de Haan et al., 2013; Pellerin et al., 2010; Nock & Ferriter, 2005). An estimated 40% of youth engaged in mental health services attend only one appointment (Saloner et al., 2014), and 50% of youth terminate by the fourth session (Harpaz-Rotem, 2004). Youth engagement in treatment activities has been shown to be a significant mediator of psychotherapeutic change (Taubner et al., 2024), and early termination from services has been associated with poorer treatment outcomes (Danko et al., 2016; Pellerin et al., 2010). As such, treatment engagement has been regarded as a critical component to bridging the gap between science and service in child and adolescent mental health services (Becker & Chorpita, 2023).

Despite decades of evidence documenting challenges with increasing and sustaining client engagement across clinical and laboratory settings, treatment engagement has historically been underemphasized in the context of evidence-based treatments (EBTs), with few interventions targeting engagement challenges directly (Becker & Chorpita, 2023). Recent work has articulated and empirically validated a multidimensional model of engagement (REACH framework; Becker et al., 2018) spanning cognitive, behavioral, and social dimensions: *Relationship* (e.g., therapeutic alliance; Shirk & Karver, 2011), *Expectancy* (e.g., beliefs about the utility of treatment; Nock & Ferriter, 2005), *Attendance* (e.g., presence at treatment sessions; Nock & Ferriter, 2005), *Clarity* (e.g., understanding about the treatment process; Shuman & Shapiro, 2002), and *Homework* (e.g., in-session participation and homework completion; Nock

& Ferriter, 2005). Although almost entirely measured as a unidimensional construct in the existing literature (e.g., attendance or homework), treatment engagement challenges are also best characterized as *multidimensional* (e.g., social, cognitive, affective, and behavioral), *dynamic* (e.g., fluctuating between and within treatment events) and *transactional* (e.g., co-constructed by the therapist, youth, and family over the course of treatment) (Becker et al., 2018; Chorpita & Becker, 2022; Haine-Schlagel & Walsh, 2015; Pullmann et al., 2013; Staudt, 2007).

Given the dynamic and transactional nature of treatment engagement (Staudt, 2007), it is important to examine how interactions between the therapist, youth, and caregivers may shape the clients' ability to engage in and benefit from services. Evidence has long documented the impact of therapist behaviors on client outcomes, highlighting the importance of identifying additional and relevant dimensions of therapist behaviors that may occur over the course of treatment delivery (King & Bickman, 2017). A recent review of observed "therapist effects" in psychotherapy hypothesized that the differential effectiveness of therapists may be attributed to specific aspects of behaviors, such as the quality of the therapist's delivery of select treatment elements (Johns et al., 2019). It is thus important to understand how therapist behaviors might influence aspects of client engagement (Karver et al., 2018). In particular, examining therapist behaviors could provide insight into best practices for improving treatment engagement and clinical outcomes.

Therapist Missteps

Therapist "missteps," or in-session therapist behaviors that are unintentional, less optimal, or noncollaborative, have been observed in diverse clinical contexts, including in the treatment of adult clients of racial, ethnic, sexual, or gender minority identities (Owen et al., 2019; Durham, 2018; Hook et al., 2016; Spengler et al., 2016; Sue et al., 2007; Mizock &

Lundquist, 2016) and specific clinical presentations, such as personality disorders (Greaves, 1998) and eating disorders (Thompson & Sherman, 1989). Therapist missteps were also found to occur across delivery of various treatment modalities, including Acceptance and Commitment Therapy (Brock et al., 2015) and psychodynamic therapy (Trimboli et al., 2016), and among mental health professionals of different training backgrounds spanning trainees, study therapists, and community providers (Knox et al., 2022; Kottler & Carlson, 2013; Park, Becker, Chu et al., 2025).

Although existing research in this area has focused predominantly on treatment with adult populations, a growing line of research has sought to measure potential therapist missteps when delivering care to youth and families and identify the context in which these missteps occur. In one study, clinical supervisors in youth community mental health settings posited that therapist missteps might vary according to features of the clinical context such as session participant type (e.g., youth-only, caregiver-only or family) and therapist preparedness to respond in the moment to engagement challenges (Park, Becker, Chu et al., 2025), suggesting that session-level and therapist-level factors may be important in understanding when missteps occur.

In close collaboration with community partners, Park and colleagues (2025) selected the term “misstep” to ensure that the term did not assume intent (i.e., that the therapist intended to engage in the observed behavior) or harm (i.e., that the behavior was definitively and universally counter-therapeutic). Akin to a moment where one is slightly off-balance, the term “misstep” was selected to emphasize the occurrence of an *accidental* error that can occur when one does not know what to do instead. This conceptualization was grounded in the well-documented clinical decision-making challenges when intervening on engagement challenges: that these missteps may occur in the absence of sufficient support or tools needed to identify and target

multidimensional engagement concerns. Missteps are thus posited as distinct from potentially harmful therapies, which refer to interventions that have been shown to have negative psychological or physical effects on clients (Lilienfeld, 2007). Missteps have been represented by a variety of terms in the current literature, and may take many forms, ranging from therapist behaviors that directly strain the therapeutic relationship such as expressing hostility or criticism (Colli et al., 2019) to the misapplication of otherwise effective strategies, such as humor (Sarink & Garcia-Montes, 2023; Valentine & Gabbard, 2014), inappropriate self-disclosure (Bottrill et al., 2010; Hanson, 2005), and advice-giving (Duan et al., 2018; Hill et al., 2023). As the field's understanding of therapist missteps deepens, understanding how and when these missteps may occur is a critical next step for the field—particularly when they occur in riskier clinical contexts, such as when low treatment engagement is present.

Challenges with Effective Management of Engagement

Effective management of low treatment engagement challenges in youth and adolescent mental health services requires a coordinated effort to (1) detect and accurately identify the presenting engagement challenges and (2) identify and deliver an evidence-based practice expected to intervene on those specific engagement challenges. The current literature suggests that both steps present distinct and well-documented challenges for therapists, particularly in the absence of adequate guidance when making clinical decisions about treatment.

Detection and Identification

Effective and timely intervention requires that one first detect that a problem exists (Youngstrom et al., 2017); however, engagement challenges have been well-documented as difficult to detect and accurately identify. Therapists have displayed an overreliance on attendance as an indicator of low engagement (Becker, Wu, et al., 2021), which has been shown

to be a late-stage indicator of disengagement (Kazdin, 1996) and does not capture engagement across other key dimensions. This pattern persists in telemental health settings: a recent study found that 79% of cases who scored at risk for low engagement went undetected by therapists, and process-oriented dimensions of engagement (i.e., relationship, expectancy, and clarity) went entirely undetected (Wu, Becker, Isenberg et al., 2026).

Even when engagement challenges are detected, accurately identifying the specific engagement challenge at hand remains difficult. This has clinical implications for treatment, as different domains of engagement respond to different types of interventions (Becker, Dickerson, et al., 2021). Therapists displayed difficulty distinguishing nuances between different engagement challenges, often utilizing terms such as “buy-in” or “parent engagement” to describe otherwise unspecified engagement challenges (Becker, Dickerson, et al., 2021). In telemental health settings, even when therapists correctly detected low engagement concerns, they were only able to accurately identify the specific client-reported REACH domains at risk less than a quarter of the time (Wu, Becker, Isenberg et al., 2026).

Selection and Delivery

The field has made meaningful progress in identifying engagement-promoting practices—the “dos” when targeting treatment engagement. Research has documented and measured discrete, well-defined therapist techniques and practices that may positively contribute to youth engagement. For example, discussing concrete barriers such as lack of transportation has been found to increase client attendance (McKay et al., 1996) and using motivational interviewing has been found to enhance readiness to participate in services (Dean et al., 2016). Common effective strategies to target the therapeutic relationship have also been identified, including validation of the youth’s feelings and experiences, actively nurturing trust, and

acknowledging and praising the efforts and engagement of the youth in therapy (Cirasola et al., 2024). A systematic review spanning four decades synthesized 30 unique clinical practice elements employed by therapists to support multidimensional engagement (Becker et al., 2018). Findings revealed that some evidence-based practices support engagement across several domains (e.g., psychoeducation was found to improve all five REACH domains of engagement), while others target select domains.

However, challenges in applying these practices have been documented when examining therapist patterns of treatment delivery. While nearly 90% of practices reported by therapists when intervening on engagement challenges were consistent with the existing evidence base, only a handful were commonly utilized by most therapists, including outreach, psychoeducation, accessibility promotion, instrumental and professional support, rapport building, and case management (Becker, Dickerson, et al., 2021). Observational coding of therapist use of engagement practices found a similar pattern of engagement practice use, in which therapists were observed to frequently use a limited set of similar practices, such as psychoeducation, rapport building, and therapist reinforcement, regardless of their training history on other EBTs or engagement practices (Wu et al., 2022). Notably, these frequently utilized practices have been shown to effectively target some domains of engagement (i.e., *Relationship*, *Attendance*, *Clarity*), but do not address others (i.e., *Expectancy*, *Homework*). Therapist training on EBTs did not significantly predict therapist knowledge of which engagement practices to apply for which engagement challenges (Reeder et al., 2024), suggesting that commonly utilized methods to support therapists in effectively delivering treatment in the presence of engagement challenges is not likely to be sufficient in increasing the repertoire and matching of practices utilized.

The consequences of this mismatch are clear. More than half of therapists reported that no practices fit their reported engagement challenges (Becker, Dickerson et al., 2021), and therapists were observed delivering practices that were not empirically justified fits for their reported challenges. Patterns such as overapplying (i.e., using an engagement practice for a challenge for which the evidence base provides no support) or underapplying practices (i.e., failing to use an engagement practice that was indicated for a presenting engagement challenge) were observed. Of the over 1,500 randomized controlled trials (RCTs) testing children’s mental health interventions, fewer than 60 tested engagement interventions (Becker et al., 2018; Lakind et al., 2022). Of those, only a subset is structured to offer the high degree of guidance that may be needed to prepare therapists in effectively intervening on these multidimensional engagement challenges. One notable exception is motivational interviewing (MI; Miller & Rollnick, 2002), a commonly utilized treatment model for treatment of adolescents presenting with high-risk concerns. MI provides detailed guidance on essential components, or “steps”, for the therapist to deliver, including reflective listening, Socratic questioning, affirming change-related statements and efforts, highlighting gaps between current behavior and desired life goals, utilizing non-confrontational responses to resistance, and collaborating on action plans (Miller & Rollnick, 2002). Most EBPs, however, do not offer comparable guidance.

Taken together, these findings suggest that there are many challenges therapists face when delivering treatment in the context of low engagement. The challenges in detecting, identifying, and targeting engagement concerns, and the subsequent gaps in therapist decision-making that can occur, may elucidate why therapist missteps may be particularly likely to occur in this context.

Therapist Preparedness and “Run Time” Decisions

One proposed hypothesis for why missteps may occur frequently in the context of low youth and family treatment engagement has posited the role of inadequate guidance for therapists to effectively identify and navigate engagement challenges as they unfold within a treatment session. In the absence of adequate guidance, therapists are often required to make “run time” decisions (see Chorpita & Daleiden, 2014), where they must respond flexibly and immediately to engagement concerns as they unfold between sessions (e.g., by consulting and adjusting treatment plans during subsequent supervision events) or within a treatment event (e.g., by deviating from the established treatment plan or intervening on a new engagement concern). Such run time decisions, in which therapists must make a *specific* clinical decision in a *particular* moment with a *particular* client and their presenting concerns, are inherently limited by the therapist’s existing knowledge available to them in that moment (e.g., training background, lived expertise, past delivery of EBTs and practices with clients of similar presentations or in similar situations). Consequently, such decisions are prone to inadvertent errors, in which therapists must do their best with limited guidance to predict what is the next best strategy (see Meehl, 1957). In the absence of guidance within most EBTs on what to do with “this here client,” therapists are left with vague boundaries separating appropriate from less appropriate variations of clinical practice (Dawes, 2005).

Several lines of evidence suggest that the specific conditions characteristic of low treatment engagement are particularly likely to produce therapist missteps. First, engagement challenges can manifest as client behaviors that have been shown to tax therapists’ attention during a treatment event, such as minimal verbal and nonverbal responses, complaints about the therapist or in-session activities, or withdrawal from planned activities (Eubanks et al., 2015;

Colli et al., 2019). Therapist negative behaviors were observed to be produced during very difficult moments in the session in response to client indicators of an alliance rupture (Colli et al., 2009). Second, provision of negative feedback from clients, particularly when expressed without a clear strategy on how to adjust the treatment approach, has been shown to elicit feelings of guilt, anxiety, incompetence, and irritation from the therapist (Coutinho et al., 2011; Hill et al., 1996; Hill et al., 2003; Moltu et al., 2010) and subsequently may make it difficult for therapists to respond effectively in the moment (Kluger & DeNisi, 1996). Third, intrapersonal challenges indicative of low engagement, including expressions of anger from the client, sporadic treatment attendance, and variability in client verbal participation in the session, have been identified as prominent antecedents to clinical errors among graduate student trainees (Knox et al., 2022). These patterns align closely with the documented challenges in managing engagement, where therapists often report not knowing what to do next to target presenting engagement challenges within the session.

As such, therapist preparedness to target these engagement challenges may be an important factor. Community experts in youth community mental health settings posited that aspects of therapist preparedness to respond in the moment to engagement challenges may impact the occurrence of therapist missteps (Park et al., 2025). However, engagement challenges may not always be transparent to therapists (Becker, Wu et al., 2021; Chorpita et al., 2025), meaning that therapists may enter a session without having identified that engagement concerns are present, nor having had adequate guidance and preparation to target those concerns. This may contribute to a clinical context in which therapists might even inadvertently behave in ways that strain engagement further. In a recent qualitative study, therapists described navigating difficult clinical situations (e.g., client appearing passive and withdrawn) as “navigating in the dark,”

where therapists felt they were unable to obtain sufficient information to make meaningful clinical decisions, which in turn, impeded their ability to access and deliver treatment (Stige et al., 2024). Importantly, what reduced the perceived difficulty of these situations was having concrete tools or interventions available and a clearer understanding of what was happening and what could be done (Stige et al., 2024)—precisely the kind of guidance that is often absent in the engagement challenge context.

This has been well-documented within the literature on therapist responses to emergent life events (ELEs), or unexpected, acute client stressors reported in treatment sessions (Chorpita, Korathu-Larson et al., 2014). Therapists who lacked adequate guidance on how to respond to these unexpected in-session stressors were more likely to deliver treatment in an unstructured manner (e.g., empathic listening, information gathering) or without using any strategies from the established treatment plan or EBT protocol—a response pattern associated with substantially slower rates of youth improvement on symptoms and functioning (Guan et al., 2017). These findings highlight how run time decisions—particularly those made without adequate guidance on when and what strategies to employ—can increase the occurrence of potential errors, and subsequently, impede effective treatment delivery.

Overall, the evidence suggests that the conditions characteristic of low treatment engagement and limited preparation and guidance in how to identify and target these multidimensional engagement challenges, may create an environment that increases the likelihood of therapist missteps occurring. Two dimensions of therapist preparedness are particularly relevant in this context. First, the recency of clinical supervision relative to treatment delivery may function as a proxy for how well-prepared a therapist is to respond to engagement challenges in the subsequent session. Research on transfer of training suggests that shorter lags

between training or supervision and subsequent performance are associated with greater retention and application of skills (Blume et al., 2010; Westman et al., 2020), implying that longer supervision-to-treatment lags may leave therapists more vulnerable to errors when engagement challenges arise. Second, the presence and quality of an established treatment plan prior to a session may serve to scaffold the cognitive burden of in-session run time decision-making. Chorpita and colleagues (2024) found that in the absence of engagement-related clinical decision-making supports, therapists were significantly less likely to receive guidance during supervision on identifying, selecting, and preparing a targeted plan for an engagement problem (Chorpita et al., 2024), further compounding concerns with detection, identification, delivery of an appropriate, engagement-promoting practice. Inadequate guidance in planning (e.g., identifying and selecting engagement practices) to target specific engagement challenges in clinical supervision may limit therapists' ability to manage additional complexities and demands that may emerge within a treatment event (e.g., youth refusal to participate in a planned session activity, in-session conflict between youth and caregiver). Treatment planning may provide therapists with a "roadmap" that reduces the likelihood of improvised, less optimal responses when unexpected engagement concerns emerge. Identifying when missteps occur and how therapist preparedness may be associated with their occurrence is thus a crucial next step to improve the quality of care for youth and families in community settings and informing how decision-making supports can reduce the occurrence of therapist missteps.

The Present Study

Efforts to delineate *when* such therapist missteps may occur over the course of treatment delivery have been stymied by notable gaps in clinical analysis and research examining broadly why clinical errors occur. A historical focus on therapist attributes (i.e., which therapists are

prone to errors), as opposed to contextual or situation attributes (i.e., what makes anyone prone to errors), has left many critical questions unanswered regarding these potential therapist missteps (see Budge, 2016). Consequently, little is known about what clinical contexts or factors may confer greater risk of therapist missteps occurring during a therapy session. Examination of potential factors associated with occurrences of therapist missteps are largely incohesive and study specific, and measurement of therapist missteps have similarly been siloed in select dimensions of engagement (e.g., relationship, attendance). Overall, the lack of a unifying taxonomy or ontology of therapist missteps have hindered systematic identification of predictors, particularly those that may more readily generalize to a plethora of treatment contexts and settings, such as therapist preparedness.

The present study aimed to address this gap directly by examining *how* and *when* therapist missteps occur in the context of low treatment engagement and examine how factors related to therapist preparedness are associated with occurrences of therapist missteps in community service settings. The study was conducted using data from the Reaching Families Study, a multisite cluster-randomized trial focused on improving use of evidence to target low treatment engagement in publicly funded youth mental health settings (Chorpita et al., 2025), one of the most common settings where youth receive mental health care (Duong et al., 2021).

Aim 1

To address the question of *how* therapists may engage in missteps, the first aim was to examine the types of missteps that occur in therapy and their relative frequencies. In line with insights from community partners related to the occurrence of therapist missteps (Park, Becker, Chu et al., 2025), frequencies of misstep occurrences were examined across key contextual factors, including session type (i.e., youth-only, caregiver-only, or family), recipient of the

misstep (i.e., youth, caregiver, or both), and the immediately preceding therapist behavior in the session (e.g., engaging in an evidence-based strategy, another misstep, or the absence of either).

Aim 2

The second aim was to examine associations between session-level predictors related to therapist preparedness (i.e., supervision-to-treatment lag, treatment plan status) and frequency of therapist misstep occurrences. Given prior research suggesting that shorter supervision-to-treatment lag were associated with greater transfer of training (Blume et al., 2010; Westman et al., 2020), it was hypothesized that longer supervision-to-treatment lags would be associated with a greater number of misstep occurrences during the therapy session (Hypothesis 1). It was also hypothesized that the presence and appropriateness of the established treatment plan would impact the frequency of misstep occurrences in the therapy session. Specifically, it was hypothesized that sessions with no established plan would show the strongest associations with misstep occurrences compared to sessions with an appropriately matched engagement treatment (Hypothesis 2).

Methods

Study Context

Data for this study came from the Reaching Families Study, a multisite cluster-randomized trial focused on promoting mental health therapists' use of evidence to engage youth and families in school mental health services (Chorpita et al., 2025). Services generally consisted of weekly individual psychotherapy sessions. Data consisted of digital recordings of psychotherapy sessions with youth and families who screened positive for the presence of at least one engagement concern. For each youth and family case enrolled, the therapist audio-recorded one to four therapy sessions ($M = 1.88$, $SD = 0.46$), yielding a total of 392 recordings. One

recording was excluded as no caregiver or client was present (i.e., therapist recorded a voicemail in which they provided an appointment reminder to the caregiver), resulting in a final analytic sample of 391 therapy sessions. Sessions conducted in Spanish were first transcribed, translated into English, then qualitatively coded. Sessions were transcribed and qualitatively coded using a behavioral observation coding system (Use of Research Evidence Content Application-Engagement; Becker & Chorpita, 2017), which included a general “misstep” code to capture intervention-agnostic therapist behaviors that occurred in the presence of engagement challenges and are potentially risky to youth and caregiver engagement.

Participants

The final sample included 391 treatment events delivered to 209 youth and families who were already receiving school mental health services and identified as experiencing low treatment engagement based on the My Thoughts About Therapy Questionnaire (MTT; Chorpita & Becker, 2022). The MTT is a 35-item, self-report measure that assesses five domains of treatment engagement as outlined by the REACH framework (Becker et al., 2018): Relationship, Expectancy, Attendance, Clarity, and Homework. Each domain consists of seven items rated on a four-point scale ranging from 0 (Strongly Disagree) to 3 (Strongly Agree), with lower scores indicating poorer engagement. Youth and caregivers included in the Reaching Families Study met criteria for low engagement (total MTT score < 14). Baseline MTT scores were used to identify the specific engagement domains in which youth and caregivers experienced challenges. Endorsement of at least one item rated “Strongly Disagree” or “Disagree” within a given subscale was used to indicate the presence of a challenge in that engagement domain. These indicators were later used to evaluate the degree of match between identified engagement challenges and the treatment plan established for each session.

Youth had a mean age of 13.08 ($SD = 3.43$) and were racially and ethnically diverse, with 43.3% identifying as Latine, 35.3% identifying as African American/Black, and 1.0% identifying as Asian American. Primary problem was diverse and reflective of youth in publicly funded mental health services, with mood-related (25.3%) occurring most frequently, followed by attention-related (19.9%), disruptive behavior-related (19%), anxiety-related (14.5%), trauma-related (11.3%) and other-related (9.5%) challenges. Services were delivered by 92 therapists ($M_{age} = 38.36$, $SD = 9.93$) who were also racially and ethnically diverse with 45.7% identifying as Latine, 39.1% identifying as African American/Black, and 3.3% identifying as Asian American. Majority of therapists were master's level clinicians (96.7%), followed by bachelor's level (2.2%) and doctoral level (1.1%) clinicians. Details of the participants have been described elsewhere (Chorpita et al., 2025).

Study Condition

Therapists in the experimental condition were provided with a coordinated knowledge system (CKS) (see Chorpita & Daleiden, 2018) and a 2-day training including use of measures to detect and prioritize multidimensional engagement concerns, review of written guides for 11 discrete clinical procedures with empirical support for engagement, and rehearsal of a one-page worksheet to support planning and monitoring in supervision and application of engagement-focused procedures in treatment events. Therapists in the enhanced care condition (i.e., Practice Guidelines condition) were alerted of each case in which either the youth or caregiver had reported an engagement concern and were provided with a set of written guides detailing the same 11 practices provided to the experimental condition resembling standard practice guides (i.e., limited to definitions and without thorough step-by-step instructions) and 1-day training

event. For Aim 2, study condition was included as a binary variable, with the CKS condition as the reference group (0).

Site

Sites included publicly funded mental health services in two geographically diverse sites with documented service inequities: a large, urban public school district on the US West Coast, and a state-run agency providing school mental health services in rural regions in the US Southeast. Both sites largely served racially minoritized youth and families. For Aim 2, site was included as a binary variable in line with analyses conducted by Chorpita and colleagues (2025), with California as the reference group (0).

Measures

Session Type

Transcripts were reviewed to determine who was present during the session. Session type was coded at the session-level as (1) youth-only, (2) caregiver-only, or (3) family sessions. Session type was not included as predictors in multilevel models and only examined descriptively in Aim 1.

Recipient of Misstep

As part of observational coding of therapist missteps by Park and colleagues (2025a, 2025b), coders assigned a code for the recipient of each misstep occurrence by reviewing each excerpt from the transcript by first examining the session type (i.e., youth-only, caregiver-only, or family). A recipient could only be coded if they were physically present in the session when the misstep occurred (i.e., a caregiver could not be coded as the recipient of the misstep in a youth-only session). For missteps occurring in family sessions, the qualitative content of the misstep was examined to determine who the therapist misstep was directed towards. Recipient of

the misstep was coded at the excerpt level. Recipient of missteps was not included as predictors in multilevel models and only examined descriptively in Aim 1.

Preceding Therapist Behavior

In the parent trial, sessions were observational coded using the Use of Research Evidence Content Application: Engagement (URECA-E; Becker & Chorpita, 2017), which assessed therapist use of evidence-based strategies related to treatment engagement. The coding system included 42 discrete codes (e.g., *Make a specific plan*, *Express empathy*) describing the therapist's behaviors. In Aim 1, step codes were utilized to characterize the therapist behavior immediately preceding each misstep occurrence for descriptive analyses only. Preceding therapist behaviors were not included as predictors in multilevel models and only examined descriptively in Aim 1.

Therapist Missteps

After treatment sessions were recorded and transcribed in the parent trial (Chorpita et. al 2025), 224 sessions (57.3%) were coded as including a possible misstep ($\kappa = .62$), spanning 979 unique instances of therapist missteps. A qualitative research design and codebook thematic analysis approach was utilized to analyze occurrences of these “missteps” in close collaboration with community informants to identify ten specific types of therapist missteps (Park, Becker, Chu et al., 2025). Coders demonstrated reliable detection of 10 specific misstep types (Fleiss' $k = .71-.88$) that also were deemed by key community partners (i.e., clinical supervisors, specialists, and program directors) with extensive years of experience at their respective study sites ($M = 16.67$, $SD = 4.73$) as potentially harmful for treatment engagement. Misstep codes and definitions are provided in Table 1. Full descriptions of the therapist misstep codebook development and qualitative coding procedures are available elsewhere (Park, Becker, Chu et al.

2025, Park, Becker, & Chorpita, 2025). For Aim 2, a session-level count variable was created to represent the total number of therapist misstep occurrences in the session (i.e., count of discrete excerpts coded as one of the ten identified misstep types).

Supervision-to-Treatment Time Lag

As part of routine care delivery, therapists engage in several clinical activities related to treatment delivery in regular supervision sessions. These activities can include reviewing presenting concerns, discussing prior delivery of treatment for a specific case, identifying specific problems to target in the next session, and making a plan for the next session. As part of the parent trial, dates of supervision and therapy sessions for each specific client were recorded. For Aim 2, supervision-to-treatment lag was operationalized as the number of days between each session and the immediately preceding supervision in which the specific client was discussed. Greater values reflected longer delays between supervision and treatment.

Treatment Plan Status

In the parent trial, the specific clinical activities that therapists engaged in during regular supervision sessions were observationally coded using the Action Cycle and Use of Evidence Behavioral Observation Coding System (ACE-BOCS; Park et al., 2020). One of the coded activities related to (1) whether a treatment plan for the next session with a specific client was made and (2) the content of the established plan (i.e., the evidence-based practice was selected as part of the treatment plan). For Aim 2, two paired codes from ACE-BOCS (i.e., *Selects Practice* and *Practice Type*) were utilized to create a categorical treatment plan status variable that reflected both the presence and appropriateness of the treatment plan established during the preceding supervision session. Based on these codes and the client's baseline engagement challenges, each session was classified and dummy coded into one of four mutually exclusive

categories: (1) No Plan, (2) Non-Engagement Plan, (3) Unmatched Engagement Plan, or (4) Matched Engagement Plan. A matched engagement plan included the presence of at least one engagement practice appropriately aligned with the client's identified engagement challenges. The levels and reference group were selected to align with prior evidence (e.g., Becker et al., 2018) outlining optimal, best practices with treatment planning when treatment engagement challenges are present.

Data Analytic Plan

Aim 1: Descriptive Frequencies of Therapist Missteps

The first aim was to establish the base rate of therapist missteps and to characterize the contexts in which misstep occurred during therapy sessions. Misstep frequencies were summarized by misstep type ($n = 10$) across key contextual factors, including session type, recipient of the misstep, and the immediately preceding therapist behavior within the session. Immediately preceding therapist behavior was classified into three mutually exclusive categories: (a) delivery of an evidence-based engagement strategy, defined as the occurrence of one of the 42 URECA-E step codes; (b) occurrence of another therapist misstep; or (c) absence of either an evidence-based step or a misstep (e.g., tangential discussion not otherwise coded).

Aim 2: Effects of Therapist Preparedness on Frequency of Therapist Missteps

The second aim examined whether session-level indicators of therapist preparedness were associated with the frequency of therapist missteps. Descriptive statistics were computed for all continuous variables in R (version 4.4.3). Correlations between all predictors and proposed covariates within each level were examined using Blimp to check for potential collinearity prior to model building. To examine whether session-level indicators of therapist preparedness were associated with the frequency of therapist missteps, a series of three-level Bayesian multilevel

models were estimated in Blimp (4.0.0) via Blimp Studio (Keller, 2025). Multilevel models were selected to account for the nested structure of the data (i.e., treatment sessions nested within clients, who were nested within therapists). Bayesian estimation was utilized as the primary estimator for the dissertation to account for the complex multilevel structure across studies, handle missing data for analyses in Study 2, and maintain analytic consistency across studies. Given that the outcome variable (i.e., number of therapist missteps occurring within a session) represented non-negative count data with substantial positive skew and overdispersion, models were specified using a negative binomial distribution, which allowed the variance of misstep counts to exceed the mean. All models included random intercepts at the client and therapist levels to account for clustering and heterogeneity in baseline misstep frequency across clients and therapists. Random slopes were not included given concerns against complex multilevel models and inclusion of numerous random effects with smaller sample sizes (van de Schoot & Miočević, 2020) and to align with prior analyses of this dataset (Chorpita et al., 2025).

Models for Aim 2 were constructed using a traditional, bottom-up building approach (Hox et al., 2017). First, an unconditional three-level negative binomial model was constructed to estimate the proportion of variance in misstep frequency attributable to clients (level 2) and therapists (level 3). The latent variable approach for intraclass correlation (ICC) calculation in count models was utilized to derive ICCs (Hox et al., 2017):

$$Variance_{Total} = Variance_{Therapist\ (Level\ 3)} + Variance_{Client\ (Level\ 2)} + \frac{\pi^2}{3}$$

Next, level 1 variables of supervision-to-treatment lag (LAG_{ijk}) and treatment plan status ($NOPLAN_{ijk}$, $NONENG_{ijk}$, $UNMATCH_{ijk}$) were added to the model. Supervision-to-treatment lag was grand mean centered to enhance interpretability. Treatment plan status was a categorical session-level predictor consisting of four mutually exclusive categories (No Plan, Non-

Engagement Plan, Unmatched Engagement Plan, Engagement Plan). This variable was dummy coded with Matched Engagement Plan as the reference category to examine whether deviations from an appropriately matched engagement plan were associated with higher rates of therapist missteps. Next, level 3 covariates of site ($SITE_k$) and study condition ($CONDITION_k$) were added to account for differences in service delivery context and provision of clinical decision-making supports. Cross-level interaction terms related to the moderating effect of study condition on the association between treatment plan status and misstep frequency were also added at this step. The model building approach is represented below, and full interpretations for all coefficients are included in Appendix B.

Level 1 Equation (Session):

$$\begin{aligned}\log(\lambda_{ijk}) = & b_{0jk} + b_{1jk}(LAG_{ijk}) + b_{2jk}(NOPLAN_{ijk}) + b_{3jk}(NONENG_{ijk}) \\ & + b_{4jk}(UNMATCH_{ijk}) \\ MISSTEP_{ijk} \sim & NegativeBinomial(\lambda_{ijk}, \phi)\end{aligned}$$

Level 2 Equation (Client):

$$\begin{aligned}b_{0jk} &= \beta_{00k} + \zeta_{0jk} \\ b_{1jk} &= \beta_{10k} \\ b_{2jk} &= \beta_{20k} \\ b_{3jk} &= \beta_{30k} \\ b_{4jk} &= \beta_{40k} \\ \zeta_{0jk} &\sim N(0, \sigma_{client}^2)\end{aligned}$$

Level 3 Equation (Therapist):

$$\begin{aligned}\beta_{00k} &= \gamma_{000} + \gamma_{001}(SITE_k) + \gamma_{002}(CONDITION_k) + v_{00k} \\ \beta_{10k} &= \gamma_{100}\end{aligned}$$

$$\beta_{20k} = \gamma_{200} + \gamma_{201}(CONDITION_k)$$

$$\beta_{30k} = \gamma_{300} + \gamma_{301}(CONDITION_k)$$

$$\beta_{40k} = \gamma_{400} + \gamma_{401}(CONDITION_k)$$

$$v_{00k} \sim N(0, \sigma_{therapist}^2)$$

Composite Equation (Full Random Intercept Model):

$$\begin{aligned} \log(\lambda_{ijk}) = & \gamma_{000} + \gamma_{001}(SITE_k) + \gamma_{002}(CONDITION_k) \\ & + \gamma_{100}(LAG_{ijk}) \\ & + [\gamma_{200} + \gamma_{201}(CONDITION_k)](NOPLAN_{ijk}) \\ & + [\gamma_{300} + \gamma_{301}(CONDITION_k)](NONENG_{ijk}) \\ & + [\gamma_{400} + \gamma_{401}(CONDITION_k)](UNMATCH_{ijk}) \\ & + v_{00k} + \zeta_{0jk} \\ MISSTEP_{ijk} \sim & NegativeBinomial(\lambda_{ijk}, \phi) \end{aligned}$$

Convergence diagnostics and model fit were examined at each model building step. In line with established guidelines (Keller & Enders, 2026), Markov Chain Monte Carlo (MCMC) burn-ins was set at 10,000 iterations and increased by increments of 10,000 until model convergence was achieved. Model convergence was defined when split-chain potential scale reduction factors (PSRFs) of less than 1.05 were reached (Du et al., 2022; Gelman et al., 2014). Additionally, post burn-in iterations were set at 10,000 and increased by intervals of 10,000 to ensure effective sample sizes for all included parameters were greater than 100 (Gelman et al., 2014). Model fit was examined at each step using the Watanabe-Akaike Information Criterion (WAIC) and Deviance Information Criterion (DIC₂; Du et al., 2024). Credible intervals (CrIs) for the exponentiated coefficients were examined to determine whether the estimated effects in the models credibly differed from one (i.e., the 95% CrI excluded one, indicating a credible

association). Posterior distribution plots were generated in R using the rblimp package (version 1.0.0; Keller, 2026) for significant effects. Syntax for multilevel models is provided in Appendix A.

Results

Aim 1: Descriptive Characterization of Therapist Missteps

On average, there were approximately 2.50 ($SD = 4.40$) misstep occurrences in each session. Relative frequencies of each misstep type are shown in the Total column of Table 2. *Advice-giving* was the most frequently occurring misstep type, accounting for 37.9% ($n = 369$) of all misstep occurrences, whereas *Eliciting shame/guilt* occurred least frequently (3.5%; $n = 34$). Of the 391 sessions, 231 (59%) sessions were youth-only, 35 (9.0%) were caregiver-only sessions, and 125 (32.0%) sessions were family sessions (i.e., at least one caregiver and youth present). Across these types of sessions, 127 (55.0%) youth-only sessions, 16 (45.7%) caregiver-only sessions, and 81 (64.8%) family sessions contained at least one misstep. Family sessions on average had the highest number of misstep occurrences per session ($M = 3.42$, $SD = 5.9$), followed by youth-only sessions ($M = 2.17$, $SD = 3.6$) and caregiver-only sessions ($M = 1.4$, $SD = 2.4$). Within each session participant type, misstep occurrences also varied by misstep type, and are displayed in the Youth-Only, Caregiver-Only, and Family columns of Table 2. In youth-only sessions, 502 total misstep occurrences spanning all 10 misstep types were coded, with *Advice-giving* occurring most frequently (37.6%; $n = 189$) and *Making a joke at the family's expense* occurring least frequently (2.0%; $n = 10$). In contrast, in caregiver-only sessions, seven misstep types accounted for the 49 misstep occurrences coded, and just five misstep types were found to have occurred more than once. The most frequent misstep type in caregiver-only sessions was also *Advice-giving* (63.3%; $n = 49$). Across family sessions, 428 total misstep

occurrences spanning all 10 misstep types were coded, with *Advice-giving* as the most frequent misstep type again (34.8%; $n = 149$). *Speculating or making assumptions* occurred least frequently in family sessions (2.6%; $n = 11$).

Frequencies of therapist missteps observed in family sessions, organized by the recipient of the misstep (i.e., youth-, caregiver-, or family-directed) are displayed in Table 3. Across the 428 misstep occurrences in family sessions, the majority were directed toward the youth (66%, $n = 284$), followed by caregiver-directed missteps (29%, $n = 124$). A smaller proportion of missteps were directed toward both the youth and caregiver simultaneously (5%, $n = 20$). Eight of the 10 types of therapist missteps occurred most frequently in interactions directed at the youth alone. In contrast, two misstep types—*Inappropriate elicitation/disclosure of information* and *Speculating or making assumptions*—were directed more frequently at caregivers than youth in family sessions.

Frequencies of therapist behaviors preceding misstep occurrences are outlined in Table 4. It is important to note that these frequencies are descriptive and do not reflect associations between therapist behaviors and subsequent missteps. Across the 979 misstep occurrences coded in the full sample, the majority (63.9%, $n = 626$) occurred immediately following the therapist's use of a structured clinical procedure, such as *express empathy*, *engage in noncontingent positives*, or *elicit perspectives about the problem*. Over one-quarter of missteps (27.5%, $n = 269$) occurred following periods in the session characterized by an absence of both a structured clinical procedure or misstep occurrences (i.e., *no codes applicable*). Consecutive misstep occurrences—instances in which one misstep immediately followed another—accounted for 10.6% of total misstep occurrences ($n = 104$).

Overall, *Advice-giving* following the use of a structured clinical procedure was the most frequently occurring scenario and accounted for 25.9% ($n = 254$) of all misstep occurrences. *Arguing or displaying doubt/disbelief* following the use of a structured clinical procedure was the next most frequent (10.5%, $n = 103$), followed by *advice-giving* in the absence of any discernable activities (8.4%, $n = 82$). *Making a joke at the family's expense* occurred most frequently in the absence of evidence-use or misstep occurrences, in contrast to the other nine misstep types, which most frequently occurred after a structured clinical procedure.

Aim 2: Effects of Therapist Preparedness on Frequency of Therapist Missteps

Descriptive statistics and correlations for the outcome variable, predictors, and covariates included in the full model are displayed in Tables 5-8. Total missteps per session ranged from 0 to 42 ($M = 2.51$, $SD = 4.43$), with a variance-to-mean ratio of 7.82 and a skewness of 3.81, indicating substantial overdispersion and right-skew. These descriptive statistics support the use of a negative binomial multilevel model. Supervision-to-treatment lag ranged from 0 to 70 days ($M = 9.00$, $SD = 10.57$). Nine sessions contained outliers for supervision-to-treatment lag (i.e., greater than three standard deviations above the mean). Sensitivity analyses were completed by comparing estimates with and without these nine sessions. Given results were substantively unchanged, the full sample was retained in primary analyses. Across sessions, the majority of sessions involved either a matched engagement plan (31.5%) or no plan (29.2%), with smaller proportions involving unmatched engagement plans (21.0%) and non-engagement plans (18.4%). Sessions were balanced across site and study condition. Correlations among session-level variables indicated that total missteps had a significant positive correlation with supervision-to-treatment lag (0.17, 95% CrI [0.07, 0.26]) and a negative correlation with maximum practice extensiveness (-0.11, 95% CrI [-0.21, -0.02]).

Model convergence diagnostics and model fit statistics for each step of the model building process are displayed in Table 9. The highest PSRFs at the final iteration were confirmed to be < 1.05 and the effective sample sizes for all estimated parameters were confirmed to be > 100 as outlined in the data analytic plan. Model fit as estimated by the conditional likelihood DIC2 and WAIC stayed relatively stable across the model building process, suggesting that the addition of level 1 and level 3 variables did not produce meaningful improvements in model fit over the null model.

The three-level negative binomial null model was first estimated in Blimp to partition the variance in therapist missteps across sessions (level 1), clients (level 2) and therapists (level 3) and is displayed on Table 10. Examination of the variance decomposition revealed significant clustering at the therapist-level. Using the latent variable approach for intraclass correlation (ICC) calculations, the proportion of variance in total misstep occurrences attributable to the therapist level was approximately 29.7% (95% CrI [0.86, 2.34]). The therapist level ICC was considerably larger than the client level ICC, which was estimated at 1.8% (95% CrI [0.01, 0.36]). Although the client level ICC was small, the client level was retained for all models because sessions were nested within clients by design. Retaining this level ensures correct estimation of standard errors and appropriately accounts for the nested structure of the data (Hox et al., 2017).

Regression coefficients and variance estimates for the full random intercept model are displayed in Table 11. The model intercept indicated an expected rate of approximately 0.52 missteps per session at the reference profile (IRR = 0.52, 95% CrI [0.28, 0.88]). Contrary to Hypothesis 1, supervision-to-treatment lag was not found to be significantly associated with frequency of therapist missteps (IRR = 1.01, 95% CrI [1.00, 1.02]). Hypothesis 2A and 2B were

also not supported by the results of this model. Compared to matched engagement plans, sessions with all other treatment plan types showed modest, but non-significant elevations in misstep rates: unmatched engagement plans (IRR = 1.76, 95% CrI [0.90, 3.39]), non-engagement plans (IRR = 1.20, 95% CrI [0.57, 2.52]), and no plan (IRR = 1.32, 95% CrI [0.71, 2.43]). However, all credible intervals included 1.0 and were therefore considered non-significant. In line, the results of the model did not find that study condition moderated the association between treatment plan status and frequency of missteps.

Site and study condition demonstrated significant associations with total misstep occurrences, and their posterior distributions are displayed in Figures 1-2. Therapists delivering treatment in South Carolina were associated with higher expected misstep rates compared to therapists in California (IRR = 1.79, 95% CrI [1.02, 3.13]). Additionally, therapists who did not receive clinical decision-making supports (i.e., Practice Guidelines condition) were associated with approximately three times the expected misstep rate compared to therapists who received the Coordinated Knowledge System (IRR = 3.18, 95% CrI [1.46, 6.82]). In the final model, significant clustering at the therapist level remained (ICC = 24.5%), while substantially less clustering remained at the client level (2.2%).

Discussion

The present study sought to explore *how* and *when* therapists may engage in a potential misstep when delivering psychotherapy services to youth and caregivers experiencing challenges with treatment engagement. Aim 1 explored *how* therapist missteps occur by examining the frequency of identified occurrences of missteps across various clinical contexts, such as session composition and type, recipient, and preceding therapist behavior in the session. Aim 2 explored

when these therapist missteps may occur by examining associations between misstep occurrences and several factors related to therapist preparedness.

How Therapists Misstep

Regarding *how* therapists may misstep, the occurrence of misstep types and their relative frequencies across various proximal factors were examined. Findings suggest that while missteps were a common occurrence across sessions, with over half of the recorded sessions containing at least one misstep occurrence, missteps represented a small portion (5.8%) of all recorded therapist behaviors within these sessions, suggesting that much remains unclear about when these missteps may happen within a session. Additionally, misstep occurrences were observed to occur in more than half of family and youth-only sessions, and just under half of caregiver-only sessions, with an average of approximately 2.5 missteps per session across session participant type. These findings are consistent with existing literature suggesting that less optimal and potentially noncollaborative therapist behaviors are a common occurrence in various treatment contexts (Eubanks et al., 2019), and present initial efforts to expand the reach of this growing body of research—which has focused predominantly on adult populations—by examining how missteps may present and unfold in youth and family treatment contexts. In addition, the occurrence of missteps across all session types (i.e., youth-only, caregiver-only, and family sessions) with notable regularity suggests that navigating engagement challenges presents universal difficulties for therapists, regardless of who is in the room.

Family sessions may pose additional complexity due to the need to maintain multiple therapeutic alliances and balance potentially conflicting perspectives and needs among family members. Closer examination of recipients of missteps in family sessions indicated that most misstep occurrences (66%) were directed towards the youth only. Notably, missteps were rarely

directed at the family (i.e., youth and the caregiver simultaneously), suggesting that missteps are more likely to be directed at only one individual at a time and may arise from challenges balancing therapist alignment with the youth and the caregiver when both are present during family sessions (Park, Becker, Chu et al., 2025).

Descriptive exploration of preceding therapist behaviors prior to each misstep occurrence provided additional context for understanding when therapist missteps occurred over the course of treatment delivery. Notably, approximately 36% of missteps occurred following another misstep (i.e., consecutive misstep occurrences) or instances where no structured clinical procedure was being implemented (i.e., *no codes applicable*). While further analyses are needed to explore whether therapists may be particularly prone to engaging in potential missteps during less structured moments of a session, this pattern descriptively aligns theoretically with prior research highlighting the additional cognitive decision-making demands present for therapists when intervening on engagement challenges (Becker, Dickerson et al., 2021). Absence of clear guidance on “what to do next,” particularly in the presence of low treatment engagement, may leave therapists with a less clear path forward between optimal (i.e., evidence-based and appropriately matched) and less optimal clinical responses.

Furthermore, clinical procedures characterized by open-ended discussions, activities or reinforcing statements to increase treatment engagement were also found to descriptively occur frequently before misstep occurrences. It is possible that the flexible and open-ended nature of these clinical procedures may create moments of increased complexity and unpredictability within the session that may tax therapists’ real-time decision-making. In line with previous research, the additional decision-making demands when intervening on engagement challenges

suggest that therapists may be more prone to misstep in less structured moments of a session; however, further analyses are needed to examine these associations.

When Therapists Misstep

Regarding *when* missteps occurred, the second aim examined associations between therapist preparedness and misstep occurrences, shedding light on key factors that may be important to consider when delivering treatment to youth and families presenting with engagement challenges.

First, supervision-to-treatment lag, operationalized by days elapsed between the therapy session and the immediately preceding supervision, showed minimal and nonsignificant associations with misstep occurrences. As such, the hypothesis that longer supervision-to-treatment lags would be associated with a greater number of misstep occurrences during the subsequent therapy session was not supported by the data. This null finding diverges from a prior publication using this dataset (Park et al., 2025), in which supervision-to-treatment lag was a marginally significant predictor of misstep occurrences. This discrepancy likely reflects methodological differences between the two analyses, as the present analyses utilized a negative binomial distribution to better accommodate the count nature and overdispersion of the outcome measure, which may have yielded a more precise estimate of this (lack of) association.

Although the timing of supervision relative to treatment delivery has been documented to significantly impact transfer of training and enhance therapist performance (Blume et al., 2010; Westman et al., 2020), the present study's operationalization of supervision-to-treatment lag may not have adequately captured relevant factors that influence supervision effectiveness and skills transfer. Whereas temporal proximity may be an important element supporting this process, other supervision elements, such as reviewing the session, clinical suggestions, and behavioral

rehearsal, may be components more critical to supporting therapist preparedness (Alfonsson et al., 2018; Choy-Brown et al., 2022; Maaß, et al., 2025) that were not captured in the present analyses. Overall, the null finding for supervision-to-treatment lag suggests that simply increasing the frequency of supervision or reducing the delay between supervision and treatment alone is unlikely to substantially reduce misstep occurrences. This finding has important implications for informing resource allocation in community mental health settings, where supervision is often a limited resource. As such, initiatives supporting the quality and content of supervision, rather than timing alone, may be needed to ensure that therapists are adequately supported for specific clinical challenges they are likely to encounter within their caseloads.

Second, associations between treatment plan status and misstep occurrences were weak and nonsignificant, suggesting that the mere presence of an appropriately matched engagement-focused treatment plan may be insufficient to guard against misstep occurrences. The present study's operationalized treatment plan status using a combined indicator representing (1) whether a practice was selected in supervision (e.g., "Let's plan to do *Goal-Setting* at the next session") and (2) the content of the selected practice (e.g., non-engagement, engagement), and (3) if an engagement practice was selected, the match between the engagement practice and the identified engagement challenges of the youth as identified by their responses on the MTT. However, this operationalization of treatment plan status may not adequately capture additional aspects of quality, specificity, and feasibility of treatment planning related to the therapist's preparedness to execute the established treatment plan. These elements, as previously discussed, may include therapist engagement in activities such as reviewing the steps involved in the engagement practice or role-playing. Prior research indicates that these supervision activities significantly impact therapist behaviors and the use of evidence-based clinical procedures in

subsequent treatment sessions (e.g., Dorsey et al., 2016; Bearman et al., 2013). As such, there may be a disconnect between treatment planning that occurs during the preceding supervision, and the execution of the established treatment plan in the therapy session. For example, the therapist may establish an appropriately matched plan to address challenges with attendance in supervision but encounter challenges with implementing this plan in the session.

Overall, these null findings suggest that although the presence and type of treatment plan established in supervision may not be significantly associated with misstep occurrences, future research may be warranted to examine a more granular indicator of treatment planning and preparation that accounts for therapist engagement in preparation activities, as well as the execution of their established treatment plan in the session. Such research could illuminate whether and under what conditions treatment planning may be associated with lower rates of therapist missteps.

Study condition emerged as credibly associated with misstep occurrences, in which therapists in the control condition were associated with an expected rate of missteps that was approximately three times higher than those in the CKS condition. This effect may be understood in light of the specific supports included in the CKS intervention, which comprehensively targeted challenges with detection and accurate identification of engagement challenges and the selection and delivery of evidence-based practices. Given the robust effects of the CKS intervention on increasing both overall and precise (i.e., matched) use of procedures supported by evidence for targeting the specific client needs across supervision and therapy sessions, as well as detailed engagement guides that provided a clear script of “steps” (e.g., therapeutic techniques to support delivery of the identified practice) (Chorpita et al., 2025), it is possible that this preparation and use of evidence may have reduced the likelihood of inadvertent missteps by

providing therapists with evidence-based options and decision-making supports to support treatment delivery. By equipping therapists with a coordinated set of decision-making supports to navigate engagement challenges, the CKS intervention may have reduced the unstructured “run time” decision-making that may create riskier conditions for missteps to occur. In this sense, the strong effect of the CKS intervention reinforces prior findings that implementing clinical decision-making “guardrails” may be a promising leverage point for reducing missteps.

Site was also credibly positively associated with misstep occurrences; however, the lower bound of credible interval of the exponentiated coefficient was close to 1.0, suggesting that this effect should be interpreted with caution. This might reflect site-level differences in EBT implementation context, including variability in organizational resources or supervision infrastructure across sites. Future analyses should consider including additional covariates that may reflect these site-level differences, such as caseload size and complexity or prior EBT training.

Limitations

Although these results indicate a significant step towards identifying specific contexts in which therapist missteps occurred when treating youth and caregivers experiencing engagement challenges, there are several limitations to consider. While frequencies of specific misstep types were coded for each treatment event, the analyses examined how key factors may be associated with the total frequency of missteps across all types, rather than examining how these factors may impact the occurrence of specific misstep types. At this point, it is unclear whether the distinction between misstep types is important; thus, future research should pursue a granular exploration of how factors of therapist preparedness may be differentially associated with the frequency of specific misstep types (e.g., demanding attention) over others (e.g., advice-giving).

Similarly, while the present study identified factors that were associated with higher occurrence of missteps, associations for other proximal factors, such as the preceding therapist's use of specific structured clinical procedures, were not examined. Although the findings provide a descriptive picture of when certain misstep types were found to occur within a treatment event (e.g., *elicit perspectives about the problem* often preceded misstep occurrences), the present study did not examine whether specific clinical procedures increased the likelihood of misstep occurrences (e.g., examining whether *elicit perspectives about the problem* confers higher risk for a misstep occurrence). Further research examining associations between the specific structured clinical procedures identified in the present study and occurrences of specific misstep types may clarify how therapists may inadvertently misstep and pivot to engage in repair behaviors and other evidence-based techniques within a session.

Overarchingly, the correlational nature of the analyses in Aim 2 precludes causal inferences. Although significant associations were observed between certain factors (e.g., study condition, site) and misstep occurrences, it cannot be concluded that these factors share a causal relationship with the observed differences in misstep rates. Additionally, it is possible that confounding variables that were not measured in the models may account for some of the observed associations. Further exploration of relevant therapist-level predictors, such as engagement in preparation activities, years of experience, training background, and caseload size, as well as session-level characteristics, such as session type, may be warranted.

Building on the present findings, analyses in Study 2 sought to examine the association between therapist missteps on youth and caregiver treatment engagement and the quality of evidence-based treatment delivery to produce actionable findings to improve therapist training and supervision in community mental health settings. Growing research in this area has outlined

the potential negative impact of therapist missteps on youth and caregiver treatment engagement and treatment delivery, which has implications for treatment gains and outcomes for families seeking services. Building on this work, Study 2 examined associations between therapist misstep occurrences on relevant outcomes, including how these missteps may be differentially associated with youth and family session-level engagement and the therapist's quality of practice delivery (e.g., extensiveness of evidence-based practices delivered within the session).

Chapter 2

What Happens in the Session: Associations Between Therapist Missteps, Delivery of Evidence-Based Care, and Session-Level Engagement

Introduction

Study 1 established that therapist missteps, or unintentional, less optimal, and noncollaborative in-session behaviors, are a measurable and prevalent phenomenon in the context of low youth and family treatment engagement, and that factors related to therapist preparedness are associated with their occurrence. A critical question that remains is whether and how the occurrence of therapist missteps is associated with key outcomes, including the quality of evidence-based practice delivery and session-level engagement of those present in the session.

Understanding “what happens” when therapist missteps occur is aligned with the broader call for research within the therapeutic alliance and therapeutic microprocesses literature. While the benefits of psychotherapy on treatment outcomes are well-established, the moment-to-moment therapist behaviors that may influence therapeutic process remain a critically understudied level of analysis (Krause, 2024). Calls for research from the psychotherapy microprocess literature posit that “every moment matters”—that the accumulation of momentary within-session interactions between the therapist and client contributes meaningfully to treatment outcomes (Luo & Levendosky, 2025). This echoes similar calls from the therapeutic alliance literature, which has long highlighted that the paucity of focus on therapist behaviors is inconsistent with dyadic models of alliance and ruptures (Kivlighan Jr et al., 2014), which stipulate that both the therapist and client co-construct these changes (Safran & Muran, 2000).

Challenges examining the association between microlevel therapeutic interventions and outcomes have been noted, given that interventions needed for a specific client at a specific moment in treatment is dynamic over the course of treatment (Krause, 2024; Crits-Cristoph & Connolly Gibbons, 2021). Nonetheless, researchers have highlighted the clinical implications of identifying specific therapist behaviors that can facilitate change and overall treatment outcomes

(Krause, 2024). Efforts to examine these associations have found that the frequency and content of therapist behaviors influence client responses at the session-level (Pereira et al., 2025).

Missteps and the Quality of Treatment Delivery

One way in which the quality of treatment delivery has been measured in publicly funded youth mental health settings is through observational coding of the extensiveness (i.e., the level of depth and detail) with which evidence-based practices (EBPs) are delivered. More extensive practice delivery reflects the degree to which a therapist delivers the associated “steps” or components comprising a given practice, and has been associated with improved client outcomes (Becker et al., 2018). Practice occurrence and extensiveness can be reliably detected and measured from session transcripts, allowing for more granular examination of service quality across levels of analysis (Wu, Becker, Chu et al., 2026).

Therapist missteps may reflect moments of disruption in treatment delivery that pose a clinically meaningful risk to the quality of practice delivery in the same session. This hypothesis is grounded in several lines of evidence examining associations between less optimal therapist behaviors and documented associations with practice delivery. First, acute and unexpected events in sessions, such as emergent life events (ELEs; Chorpita, Korathu-Larson et al., 2014), have been associated with increased therapist use of unstructured or non-evidence-based strategies, a pattern linked with substantially slower rates of youth improvement (Guan et al., 2019). In the absence of adequate guidance on how to respond to these unexpected moments and return to an established treatment plan or respond flexibly in an evidence-based manner, therapists may face an increased risk of either ignoring the concern and moving forward, or improvising a response that deviates from what has already been delivered without a structured

alternative (Chorpita & Daleiden, 2014). Findings suggest that neither option is ideal and may lead to a degradation in clinical care.

Second, research on therapist response to challenging client behaviors suggest that difficult clinical moments more broadly are associated with less organized and more variable therapist behaviors. When therapists were uncertain about how to respond to problematic client behaviors (e.g., disengagement, resistance), their in-session behaviors were more variable and harder to characterize qualitatively (Pereira et al., 2025). In contrast, therapist behaviors in response to client progress were more organized, evidence-based, and easily recognizable (Pereira et al., 2025). Third, qualitative research on therapists' experiences following clinical errors suggests that such missteps may elicit difficult internal states that may limit the therapist's ability to act in the best interest of the client in the session. Therapists noted experiencing negative emotional responses, self-doubt, feelings of inadequacy, and concerns regarding the client's response and future engagement in services (Knox et al., 2022; Casement, 2002; Sharma & Fowler, 2016; Waehler & Grandy, 2016).

While the direct association between therapist missteps and extensiveness of practice delivery have not been directly examined, it is hypothesized that therapist misstep occurrences may represent similar "in the moment" disruptions in service delivery. It is well established that addressing treatment engagement challenges presents distinct challenges for clinical decision-making, including accurate assessment of presenting engagement challenges and informed selection and delivery of engagement-promoting EBPs. Following inadvertent missteps, therapists may have difficulty returning to an established treatment plan or deviating flexibly and effectively to meet an imminent engagement concern, reducing the overall extensiveness and quality of practices delivered in the session. The relationship between practical guidance and

practice extensiveness is illustrated in findings from Wu and colleagues (2022). Therapists who used a specific protocol to guide a session delivered practices with higher extensiveness, suggesting that this guidance may have increased provider awareness of specific techniques to utilize and in turn scaffolded more thorough practice delivery. In the absence of a guiding protocol, therapists delivered practices with lower breadth and depth (Wu et al., 2022), a pattern consistent with the broader evidence that lower supports are associated with less organized clinical behavior (Pereira et al., 2025; Guan et al., 2017).

Missteps and Associations with Engagement Outcomes

Preliminary evidence suggests that therapist missteps may be meaningfully associated with treatment engagement, both within and across treatment sessions. Therapist use of “negative interventions,” including unwanted advice-giving, expressing hostility or sarcasm, criticism, and excessively insisting on a selected topic, were associated with greater client indicators of ruptures in the therapeutic alliance, such as strong disagreement or withdrawal (Colli et al., 2009; 2019). Therapist behaviors indicative of lower responsiveness, such as forgetting what clients had previously said, not acknowledging client emotions, or pushing the youth to talk, were found to negatively impact alliance (Creed & Kendall, 2005; Russell et al., 2008) and was associated with early termination from services (Elkin et al., 2014). The extent to which therapists caused or exacerbated within-session alliance ruptures predicted client drop out (Eubanks et al., 2019). Several therapist-reported behaviors, including failing to recognize or address a problem in treatment, advice-giving, and allowing personal feelings about a client to interfere with treatment, were also found to be associated with premature termination (Piselli et al., 2011). Qualitative analyses of sessions with adolescents who had prematurely terminated services and reported dissatisfaction with treatment similarly identified the occurrence of therapist behaviors

that contributed to alliance ruptures, including minimal responses, rigid adherence to an activity or session agenda, and prioritization of risk concerns over other clinical needs (O’Keeffe et al., 2020). A systematic review found that therapist in-session behaviors were shown to both positively and negatively influence therapeutic alliance in youth mental health settings (Ryan et al., 2023). Whereas behaviors such as collaboration (Creed & Kendall, 2005), exploration of feelings (Karver et al., 2008; Russell et al., 2008), and assessment of motivation (Jungbluth & Shirk, 2009) were positively associated with alliance, distortion of information (Karver et al., 2008) and criticism (Russell et al., 2008) negatively impacted alliance.

Collectively, these findings suggest that therapist misstep-like behaviors are associated with decreases in engagement, particularly straining the therapeutic relationship and attendance. Although therapist missteps may most visibly manifest as ruptures in the therapeutic relationship or premature termination from services, it is hypothesized that the consequences of therapist missteps likely extend to other dimensions of engagement, including expectancy, clarity, and homework (Park, Becker, Chu et al., 2025). Adverse impacts of misstep-like behaviors are diverse and include interference with client progress (Henry et al., 1990); clients feeling unheard and unhelped by the therapist (Knox et al., 2011), decreased openness (Park, Becker, Chu et al., 2025); and abrupt or problematic termination from services (Renk & Dinger, 2002; Knox et al., 2010; Park, Becker, Chu et al., 2025).

However, the existing research has had notable limitations that have hindered a thorough examination of the relationship between therapist missteps and youth and family treatment engagement. Few studies have explored the effect of missteps on all five domains of treatment engagement (i.e., REACH; Becker et al., 2018). Additionally, most therapist missteps have been measured in response to specific precipitating client behaviors (e.g., Colli et al., 2009), making it

challenging to ascertain the directionality of these associations—for example, whether the therapist behavior drove the change in session engagement or whether the client’s prior low engagement elicited the misstep. Park and colleagues (2025) sought to address noted measurement limitations by anchoring the detection and measurement of therapist missteps to specific therapist behaviors across the session, regardless of how youth, caregivers, and therapists subsequently responded. Ten misstep types were identified with excellent reliability and preliminary face validity (Park, Becker, Chu et al., 2025); however, associations between missteps, session engagement, and quality of treatment delivery have yet to be examined.

The Present Study

Efforts to examine associations between therapist missteps, quality of treatment delivery, and engagement have been limited to date. The present study addresses these gaps by examining whether the frequency of misstep occurrences within a session are associated with (1) the quality of evidence-based practice delivery in that session, and (2) session-level engagement of the youth and caregivers present. Elucidating these associations may inform best practices in supporting therapists in employing effective resolution strategies in the moment.

Aim 1

The first aim was to examine the association between the frequency of misstep occurrences and the quality of evidence-based practice delivery within a session, and to examine whether the presence of an established treatment plan moderated this relationship. Quality of evidence-based practice delivery was measured by the average practice extensiveness rating achieved across all practices delivered within the session. Consistent with findings that therapists are more likely to utilize more unstructured and non-evidence-based strategies when unexpected events occur in a session (e.g., Guan et al., 2019; Pereira et al., 2025), it was hypothesized that

greater frequency of misstep occurrences would be associated with a lower average extensiveness rating for the session (Hypothesis 1). Consistent with evidence that treatment planning activities support therapist use of evidence during treatment (e.g., Dorsey et al., 2015; Scott & Lewis, 2015); it was expected that the presence of an established treatment plan prior to the session would moderate the relationship between misstep frequency and average practice extensiveness (Hypothesis 2). Associations between misstep frequency and the maximum practice extensiveness rating achieved for any single practice delivered in the session were also examined as a secondary, exploratory analysis.

Aim 2

The second aim was to examine the association between the frequency of therapist misstep occurrences and session-level engagement of youth and caregivers. For each session, session-level engagement responses were observationally coded for indicators of lower or higher engagement across the five REACH domains, and response scores were summed across domains to create an aggregate score reflecting overall engagement for that session. Session-level engagement responses were recorded for each participant present in the session (i.e., youth, caregiver, or both). Youth and caregiver engagement were examined separately given expected differences in how therapist-level and client-level factors may differentially contribute to engagement across these groups.

Consistent with preliminary findings linking therapist misstep-like behaviors with decreases in specific domains of engagement (e.g., Piselli et al., 2011; O’Keeffe et al., 2020; Eubanks et al., 2019) and qualitative findings from community providers highlighting the potential impact of therapist missteps on engagement (Park, Becker, Chu et al., 2025), it was

hypothesized that greater frequency of misstep occurrences would be associated with lower youth and caregiver overall engagement scores at each session (Hypothesis 3).

Methods

Study Context and Participants

The present study used the same sample of youth, caregivers, therapists, and treatment sessions described in Study 1. Data for this study came from the Reaching Families Study, a multisite cluster-randomized trial conducted in publicly funded mental health services (see Chorpita et al., 2025). Full details regarding study context, participant characteristics, data collection procedures, and observational coding are provided previously in Study 1. Site and study condition utilized the same coding schema as Study 1.

The data set included 391 audio-recorded psychotherapy sessions delivered to 209 youth and families identified as experiencing low treatment engagement. Sessions varied by audience (i.e., who was present for the sessions), with 231 (59%) youth-only sessions, 35 (9%) caregiver-only sessions, and 125 (32%) family sessions. Analyses for Aim 1 utilized the full sample of 391 sessions. Analyses for Aim 2 utilized two overlapping subsamples of the 391 sessions—a sample of sessions where the youth was present (i.e., youth-only and family sessions; $n = 356$, 91.0%) and a sample of sessions where a caregiver was present (i.e., caregiver-only and family sessions; $n = 160$, 40.9%). These subsamples were utilized for the analyses for Aim 2 to conduct parallel multilevel analyses with different audience-specific dependent variables. One set of models examined youth session-level engagement as the dependent variable, while the second set of models examined caregiver session-level engagement.

Measures

Therapist Missteps

Therapist missteps were observationally coded as outlined in Park and colleagues (2025) and represented occurrences of inadvertent therapist behaviors occurring in the context of engagement challenges. As used in Study 1, a session-level count variable was utilized to represent the total number of misstep occurrences within a session.

Average Extensiveness of Practice Delivery (Hypothesis 1)

In the Reaching Families trial, each session was observationally coded for how extensively (e.g., breadth, quality, depth) the therapist delivered evidence-based practices. All practices delivered within a session were assigned an extensiveness rating ranging 1-5, with higher ratings indicating greater extensiveness. To obtain a metric of the overall quality of practice delivery for a given treatment event, the average of the extensiveness ratings assigned to all practices delivered in the treatment event was obtained for each session. Consistent with findings that therapists are more likely to utilize more unstructured and non-evidence-based strategies when unexpected events occur in a session (e.g., Guan et al., 2019; Guan et al., 2017), it was hypothesized that the frequency of misstep occurrences would be associated with a lower average extensiveness rating for all practices delivered within the session.

Maximum Extensiveness of Practice Delivery (Exploratory Analyses)

In follow-up exploratory analyses, the maximum extensiveness of practice delivery was utilized to examine associations with total misstep occurrences and session-level engagement. The maximum practice extensiveness rating refers to the therapist's "best" performance across all practices delivered within a session and was measured as a session-level variable capturing the highest practice extensiveness rating achieved for a session. Maximum practice extensiveness

rating was measured on the same 1 to 5 scale (1 = minimal extensiveness, 5 = high extensiveness).

Number of evidence-based practices delivered per session

The number of evidence-based practices delivered within each session was observationally coded as part of the larger trial. In the present study, the total number of practices delivered was measured at the session-level and added as a covariate (exploratory addition to originally proposed analyses) to reduce confounding the occurrence of missteps with “lighter” sessions that consisted of fewer practices delivered within the session.

Presence of Treatment Plan

The presence or absence of a treatment plan was measured utilizing a binary indicator. Using similar methods to Study 1, supervision sessions were observationally coded for the occurrence of a code *Selects Practice*, indicating whether an evidence-based practice was selected for delivery in the next session (Park et al., 2020). If the code *Selects Practice* occurred at least once in the preceding supervision, the subsequent session was characterized as Plan Present (0). If the code *Selects Practice* did not occur, the session was characterized as No Plan (1). Consistent with findings suggesting the impact of treatment planning activities on therapist behaviors and use of evidence during a treatment event, it was hypothesized that the presence of an established treatment plan prior to the treatment event would significantly moderate the relationship between the frequency of therapist missteps and the average practice extensiveness achieved for the session.

Session-level youth and caregiver engagement (Hypothesis 3)

Session-level engagement of the audience present in the session (i.e., youth, caregiver) was observationally coded as part of the Reaching Families Study. Session-level engagement

scores were only assigned for individuals present in the session (e.g., no caregiver scores for youth-only sessions). The session-level engagement score represented a sum of scores across five domains reflecting the REACH engagement framework (i.e., Relationship, Expectancy, Attendance, Clarity, Homework). For each individual present in the session, a session-level engagement scores for each of the five domains was scored on a scale of 1 to 5. Scores of 1 indicated low treatment engagement for each specified domain (e.g., for *Attendance*, explicit evidence of being late to session). Scores of 5 indicated high engagement for each domain (e.g., for *Attendance*, explicit evidence of being on time or early to session). Session-level engagement scores for each domain were recoded to a 0 to 4 coding schema to allow for a meaningful interpretation of the null value and summed to create a composite youth and/or caregiver engagement score for each session.

Baseline youth and caregiver engagement

Baseline youth and caregiver engagement were measured using the youth and caregiver versions of the My Thoughts About Therapy Questionnaire (MTT; Chorpita & Becker, 2022). The youth and caregiver versions of the MTT are 35-item, self-report measures that assess five domains of treatment engagement as outlined by the REACH framework (Becker et al., 2018). Each domain consists of seven items rated on a four-point scale ranging from 0 (Strongly Disagree) to 3 (Strongly Agree), with lower scores indicating poorer engagement. Items on the MTT assessed for the presence of domain-specific engagement challenges. For example, within *Relationship*, items included “I like meeting with my counselor” and “My counselor understands my culture and values,” whereas items within *Expectancy* included “I think my counselor can help me” and “I believe the work I do with my counselor will help me.” Baseline engagement challenges for the youth and caregiver were operationalized as the total number of items on the

MTT-Y and MTT-CG endorsed as “Strongly Disagree” or “Disagree.” Greater number of items endorsed on the MTT represented the presence of more engagement concerns across domains.

Data Analytic Plan

Descriptive statistics and correlations between predictors and covariates were examined for possible collinearity prior to model building. Three-level Bayesian multilevel models were estimated in Blimp (4.0.0) via Blimp Studio (Keller, 2025). Multilevel models were selected to account for the nested structure of the data (i.e., treatment sessions nested within clients, who were nested within therapists). In line with Study 1, all models included random intercepts at the client and therapist levels to account for clustering, and random slopes were not included. Models for Aim 1 and 2 were constructed using a traditional, bottom-up building approach (Hox et al., 2017) used in Study 1.

Convergence diagnostics and model fit were examined at each model building step. In line with established guidelines (Keller & Enders, 2026), Markov Chain Monte Carlo (MCMC) burn-in periods were run with 10,000 iterations and increased by intervals of 10,000 until the model converged. Model convergence was defined when split-chain potential scale reduction factors (PSRFs) of less than 1.05 were reached (Du et al., 2022; Gelman et al., 2014). Additionally, post burn-in iterations were set at 10,000 and increased by intervals of 10,000 to ensure effective sample sizes for all included parameters were greater than 100 (Gelman et al., 2014). Model fit was examined at each step using the marginal Watanabe-Akaike Information Criterion (WAIC) and Deviance Information Criterion (DIC₂; Du et al., 2024). Credible intervals (CrIs) were examined to determine whether the estimated effects in the models significantly differed from zero (i.e., excluded zero within the 95% CrI). Posterior distribution plots were

generated in R using the rblimp package (version 1.0.0, Keller, 2026) for significant effects.

Syntax for multilevel models is provided in Appendix A.

Aim 1: Effect of Therapist Missteps on Average Practice Quality

First, an unconditional three-level model was constructed to estimate the proportion of variance in average practice extensiveness attributable to sessions (level 1), clients (level 2) and therapists (level 3). Next, level 1 variables of total misstep occurrences ($Misstep_{ijk}$), total number of practices delivered in the session ($TotPractices_{ijk}$), and plan status ($Plan_{ijk}$) were added to the model. Total number of practices delivered in a session was grand mean centered to enhance interpretability. Next, level 3 covariates of site ($SITE_k$) and study condition ($CONDITION_k$) were added to account for differences in service delivery context and provision of clinical decision-making supports. At this step, an interaction term was also added to examine whether the presence or absence of a treatment plan for the session moderated the association between misstep occurrences and the session-level average practice extensiveness rating for the session. The model building approach is represented below, and full interpretations for all coefficients are included in Appendix B.

Level 1 Equation (Session)

$$\begin{aligned} AvgExt_{ijk} = & b_{0jk} + b_{1jk}(Misstep_{ijk}) + b_{2jk}(TotPractices_{ijk}) + b_{3jk}(Plan_{ijk}) \\ & + b_{4jk}(Misstep_{ijk} \times Plan_{ijk}) + \varepsilon_{ijk} \\ \varepsilon_{ijk} \sim & iid N(0, \sigma_\varepsilon^2) \end{aligned}$$

Level 2 Equation (Client)

$$\begin{aligned} b_{0jk} = & \beta_{00k} + \zeta_{0jk} \\ b_{1jk} = & \beta_{10k} \\ b_{2jk} = & \beta_{20k} \end{aligned}$$

$$b_{3jk} = \beta_{30k}$$

$$b_{4jk} = \beta_{40k}$$

$$\zeta_{0jk} \sim N(0, \sigma_{client}^2)$$

Level 3 Equation (Therapist)

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(Site_k) + \gamma_{002}(StudyCondition_k) + v_{00k}$$

$$\beta_{10k} = \gamma_{100}$$

$$\beta_{20k} = \gamma_{200}$$

$$\beta_{30k} = \gamma_{300}$$

$$\beta_{40k} = \gamma_{400}$$

$$v_{00k} \sim N(0, \sigma_{therapist}^2)$$

Composite Equation (Full Random Intercept Model):

$$\begin{aligned} PRAC_{EXT_{ijk}} = & \gamma_{000} + \gamma_{001}(Site_k) + \gamma_{002}(StudyCondition_k) \\ & + \gamma_{100}(Misstep_{ijk}) \\ & + \gamma_{200}(TotPractices_{ijk}) \\ & + \gamma_{300}(Plan_{ijk}) \\ & + \gamma_{400}(Misstep_{ijk} \times Plan_{ijk}) \\ & + v_{00k} + \zeta_{0jk} + \varepsilon_{ijk} \end{aligned}$$

Exploratory Analyses: Effect of Maximum Practice Quality on Therapist Missteps

Exploratory analyses were motivated in part by an unexpected null finding for study condition in the primary Aim 1 analyses (see Originally Proposed Models and Deviations). Therapist missteps occurred alongside practice delivery, and the present analyses are unable to determine directionality of the relationship between misstep occurrences and indicators of the

quality of practice delivery. Additionally, the aggregate nature of the average practice extensiveness outcome in the originally proposed model may obscure meaningful within-session variation in the quality of practice delivery across practices occurring within the session. One possibility is that the condition effects on practice delivery may be better captured by a non-aggregate indicator that reflects the therapist's peak performance within a session rather than average performance across practices. Theoretically, sessions that are characterized by the occurrence of at least one practice with high breadth and depth may reflect higher therapist preparedness for the session (i.e., a therapist who knew what to do and how to do it well). From this perspective, the maximum practice extensiveness achieved across all practices delivered within a session (i.e., the therapist's "best" performance) can be conceptualized as an indicator of therapist preparedness for that session, which may both reflect and constrain the occurrence of missteps, which are theorized to occur when the therapist is less clear about what to do next. Given the bidirectional theoretical plausibility of this relationship and the inability of the present study design to establish directionality, exploratory follow-up analyses were conducted to examine the association between the maximum practice extensiveness for the session and misstep occurrences, with maximum practice extensiveness included as a session-level variable in the Study 1 Aim 2's primary model structure. Maximum practice extensiveness was grand mean centered to enhance interpretability. Exploratory analyses were conducted following the completion of the originally proposed Aim 1 analyses.

Level 1 Equation (Session):

$$\begin{aligned} \log(\lambda_{ijk}) = & b_{0jk} + b_{1jk}(LAG_{ijk}) + b_{2jk}(NOPLAN_{ijk}) + b_{3jk}(NONENG_{ijk}) \\ & + b_{4jk}(UNMATCH_{ijk}) + b_{5jk}(MaxExtensiveness_{ijk}) \\ MISSTEP_{ijk} \sim & NegativeBinomial(\lambda_{ijk}, \phi) \end{aligned}$$

Level 2 Equation (Client):

$$b_{0jk} = \beta_{00k} + \zeta_{0jk}$$

$$b_{1jk} = \beta_{10k}$$

$$b_{2jk} = \beta_{20k}$$

$$b_{3jk} = \beta_{30k}$$

$$b_{4jk} = \beta_{40k}$$

$$b_{5jk} = \beta_{50k}$$

$$\zeta_{0jk} \sim N(0, \sigma_{client}^2)$$

Level 3 Equation (Therapist):

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(SITE_k) + \gamma_{002}(CONDITION_k) + v_{00k}$$

$$\beta_{10k} = \gamma_{100}$$

$$\beta_{20k} = \gamma_{200} + \gamma_{201}(CONDITION_k)$$

$$\beta_{30k} = \gamma_{300} + \gamma_{301}(CONDITION_k)$$

$$\beta_{40k} = \gamma_{400} + \gamma_{401}(CONDITION_k)$$

$$\beta_{50k} = \gamma_{500}$$

$$v_{00k} \sim N(0, \sigma_{therapist}^2)$$

Composite Equation (Full Random Intercept Model):

$$\log(\lambda_{ijk}) = \gamma_{000} + \gamma_{001}(SITE_k) + \gamma_{002}(CONDITION_k)$$

$$+ \gamma_{100}(LAG_{ijk})$$

$$+ [\gamma_{200} + \gamma_{201}(CONDITION_k)](NOPLAN_{ijk})$$

$$+ [\gamma_{300} + \gamma_{301}(CONDITION_k)](NONENG_{ijk})$$

$$+ [\gamma_{400} + \gamma_{401}(CONDITION_k)](UNMATCH_{ijk})$$

$$+\gamma_{500}(MaxExtensiveness_{ijk})$$

$$+ v_{00k} + \zeta_{0jk}$$

$$MISSTEP_{ijk} \sim NegativeBinomial(\lambda_{ijk}, \phi)$$

Aim 2: Effect of Therapist Missteps on Youth and Caregiver Session Engagement

Findings from the Aim 1 exploratory analyses informed a revision to the originally proposed Aim 2 models. Specifically, given the theoretical and empirical evidence suggesting that maximum practice extensiveness reflects session-level treatment quality that co-occurs with misstep occurrences, maximum practice extensiveness was added as a session-level covariate in the Aim 2 models. This addition was motivated by theoretical interest in isolating the unique association between misstep occurrences and engagement, above and beyond general session-level treatment quality. This represents a deviation from the originally proposed analyses that is justified on both theoretical and empirical grounds and is described further in the Originally Proposed Models and Deviations section below.

First, unconditional three-level models were constructed to estimate the proportion of variance for youth session-level engagement (Aim 2a) and caregiver session-level engagement (Aim 2b) attributable to sessions (level 1), clients (level 2) and therapists (level 3). Next, level 1 variables of total misstep occurrences ($Misstep_{ijk}$) and maximum practice extensiveness for the session ($MaxExtensiveness_{ijk}$) were added to the model. Then, a level 2 variable of baseline youth or caregiver engagement challenges as measured by the MTT-Y or MTT-CG was added. Lastly, level 3 covariates of site ($SITE_k$) and study condition ($CONDITION_k$) were added. Maximum practice extensiveness and youth and caregiver baseline engagement challenges were grand mean centered to enhance interpretability. The model building approach for Aim 2 is represented below, and full interpretations for all coefficients are included in Appendix B.

Level 1 Equation (Session):

$$Y_{ijk} = b_{0jk} + b_{1jk}(\text{Missteps}_{ijk}) + b_{2jk}(\text{MaxExtensiveness}_{ijk}) + \varepsilon_{ijk}$$

$$\varepsilon_{ijk} \sim iid N(0, \sigma_{\varepsilon}^2)$$

Level 2 Equation (Client):

$$b_{0jk} = \beta_{00k} + \beta_{01k}(\text{BaselineMTT}_{jk}) + \zeta_{0jk}$$

$$b_{1jk} = \beta_{10k}$$

$$b_{2jk} = \beta_{20k}$$

$$\zeta_{0jk} \sim N(0, \sigma_{client}^2)$$

Level 3 Equation (Therapist):

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{Site}_j) + \gamma_{002}(\text{StudyCondition}_k) + v_{00k}$$

$$\beta_{01k} = \gamma_{010}$$

$$\beta_{10k} = \gamma_{100}$$

$$\beta_{20k} = \gamma_{200}$$

$$v_{00k} \sim N(0, \sigma_{therapist}^2)$$

Composite Equation (Full Random Intercept Model):

$$Y_{ijk} = \gamma_{000} + \gamma_{001}(\text{Site}_j) + \gamma_{002}(\text{StudyCondition}_k)$$

$$+ \gamma_{010}(\text{BaselineMTT}_{jk})$$

$$+ \gamma_{100}(\text{Missteps}_{ijk})$$

$$+ \gamma_{200}(\text{MaxExtensiveness}_{ijk})$$

$$+ v_{00k} + \zeta_{0jk} + \varepsilon_{ijk}$$

Originally Proposed Models and Deviations

The analyses presented in this study differed slightly from the models originally proposed. Equations for the originally proposed analyses are provided in Appendix C. First, for Aim 1, total number of practices delivered per session was added as a session-level covariate prior to model estimation. Average practice extensiveness for the session is computed across all practices delivered within a session, meaning sessions with fewer practices delivered would have an average extensiveness rating that is more heavily influenced by any single practice rating. The total number of practices delivered was included as a covariate and allow the association between missteps and average practice extensiveness to be interpreted independently of session-level variation in the number of practices delivered.

Secondly, the primary Aim 1 analyses yielded an unexpected null finding for the effect of study condition on average practice extensiveness. This finding was theoretically puzzling and potentially inconsistent with prior analyses on this dataset by Chorpita and colleagues (2025), which found that therapists in the CKS condition displayed odds ratios of 2.89 for delivering any practice and 11.96 for delivering a matched, engagement-focused practice in the therapy session over therapists in the PG condition. Although there were key differences between analyses in the present study and Chorpita et al. (2025)—specifically, the present study examined the extensiveness of practices delivered, rather than the odds that the practice occurred in the session—this finding conflicted with established evidence and warranted closer examination. Given the theoretical and bidirectional plausibility of the relationship between indicators of session-level treatment quality and misstep occurrences, and the limitations of utilizing an aggregate measure of treatment quality, a non-aggregate indicator (i.e., maximum practice extensiveness achieved within the session) was added as a session-level predictor utilizing the

Study 1 Aim 2 model structure. Maximum practice extensiveness was treated as a co-occurring session-level variable rather than a causal antecedent or consequence of misstep occurrences.

Lastly, based on the findings of the exploratory analyses, which demonstrated a credible negative association between maximum practice extensiveness and misstep occurrences, maximum practice extensiveness was added as a session-level covariate in Aim 2 models prior to model estimation. This addition served to isolate the unique association between misstep occurrences and session-level engagement, above and beyond general session-level treatment quality. Consistent with the nondirectional framing applied in the exploratory analyses, maximum practice extensiveness was treated as a co-occurring session-level variable in Aim 2, and directionality among session-level variables could not be established from the present study design.

Results

Aim 1: Effect of Therapist Missteps on Average Practice Quality

Descriptive statistics and correlations for the outcome variable, predictors, and covariates included in the full model are displayed in Tables 12-14. Average practice extensiveness ranged from 0 to 5 ($M = 1.72$, $SD = 0.65$), with higher scores reflecting more thorough delivery of evidence-based practices. Three sessions received an average practice extensiveness rating of zero because no identifiable practice was delivered in the session. The distribution of average practice extensiveness clustered between extensiveness ratings of 1.0 and 2.0, reflecting average delivery with minimal to moderate breadth. Total practices delivered per session ranged from 0 to 11 ($M = 4.23$, $SD = 2.07$), with a roughly normal distribution centered around 3 to 5 practices per session. Correlations between session-level study variables showed that average practice extensiveness had a small negative correlation with total misstep occurrences (-0.13 , 95% CrI [-

0.22, -0.03]) and a small positive correlation with total number of practices delivered (0.16, 95% CrI [0.07, 0.26]).

Model convergence diagnostics and model fit statistics for each step of the model building process is displayed in Table 15. The highest PSRFs at the final iteration were confirmed to be < 1.05 and the effective sample sizes for all estimated parameters were confirmed to be > 100 as outlined in the data analytic plan. Model fit as estimated by the conditional likelihood DIC2 and WAIC showed slight improvement, suggesting small improvements in model fit over the null model.

The unconditional model was estimated to partition the variance in average practice extensiveness across sessions (level 1), clients (level 2), and therapists (level 3) and is displayed in Table 16. Approximately 15.1% of the variance in average practice extensiveness was attributable to therapist-level nesting (95% CrI [4.9%, 28.0%], and 8.9% was attributable to client-level nesting (95% CrI [1.0%, 24.0%]). Substantial variation was at the session-level (73.8%, 95% CrI [60.4%, 85.8%]).

Regression coefficients and variance estimates for the full random intercept model are displayed in Table 17. The model intercept indicated an expected average practice extensiveness of 1.87 for the session at the reference profile ($b = 1.87$, 95% CrI [1.73, 2.01]). Findings supported Hypothesis 1: total misstep occurrences were credibly associated with average practice extensiveness for the session ($b = -0.02$, 95% CrI [-0.04, -0.01]). Hypothesis 2 was also not supported. The presence of a treatment plan did not appear to moderate the association between total misstep occurrences and average practice extensiveness for the session ($b = 0.004$, 95% CrI [-0.03, 0.04]). These findings suggest that the association between misstep occurrences and

average practice extensiveness did not differ significantly based on whether a treatment plan for the session was present.

Exploratory Analyses: Effect of Maximum Practice Quality on Therapist Missteps

Model convergence diagnostics and model fit statistics are displayed in Table 18, and the full random intercepts model is displayed in Table 19. Maximum practice extensiveness was credibly associated with misstep occurrences (IRR = 0.81, 95% CrI [0.70, 0.96]). Specifically, each one-point increase in the maximum practice extensiveness rating for the session was associated with an expected rate of misstep occurrences that was approximately 19% lower, accounting for treatment plan status, supervision-to-treatment lag, site, and study condition. Additionally, study condition remained credibly associated with misstep occurrences following the addition of maximum practice extensiveness into the full model (IRR = 3.01, 95% CrI [1.39, 6.65]). Specifically, therapists in the Practice Guidelines condition—who did not receive coordinated clinical decision-making supports to target engagement challenges—were associated with an expected rate of misstep occurrences that was approximately three times the expected rate for therapists in the Coordinated Knowledge System (CKS) condition, holding all other predictors and covariates constant. Posterior distributions for maximum practice extensiveness and study condition are displayed in Figures 3-4.

Aim 2: Effect of Therapist Missteps on Youth and Caregiver Session Engagement

Descriptive statistics and correlations for the outcome variable, predictors, and covariates included in the youth and caregiver full models are displayed in Tables 20-23. Youth session-level engagement scores ranged from 1 to 19 ($M = 10.16$, $SD = 2.80$) and caregiver scores ranged from 3 to 18 ($M = 2.45$, $SD = 2.45$), with higher scores reflecting more higher overall session-level engagement. Maximum practice extensiveness in youth-present sessions had a mean of

2.60 ($SD = 1.16$) and 2.48 ($SD = 1.20$) in caregiver-present sessions. Number of at-risk items endorsed on the baseline MTT ranged from 1 to 20 for youth-present sessions ($M = 4.89$ $SD = 4.92$) and 1 to 14 for caregiver-present sessions ($M = 2.94$, $SD = 5.67$).

Correlations between session-level study variables showed that youth session-level engagement had a negative correlation with total misstep occurrences (-0.22, 95% CrI [-0.32, -0.12]) and a positive correlation with the maximum practice extensiveness (0.43, 95% CrI [0.34, 0.51]). A similar pattern emerged for caregiver session-level engagement, which was negatively correlated with total misstep occurrences (-0.28, 95% CrI [-0.42, -0.13]) and positively correlated with maximum practice extensiveness (0.43, 95% CrI [0.19, 0.48]). Total missteps were negatively associated with maximum practice extensiveness in youth-present (-0.14, 95% CrI [-0.24, -0.04]) and caregiver-present sessions (-0.04, 95% CrI [-0.20, 0.12]).

Model convergence diagnostics and model fit statistics for both models are displayed in Table 24-25. The highest PSRFs at the final iteration were confirmed to be < 1.05 and the effective sample sizes for all estimated parameters were confirmed to be > 100 as outlined in the data analytic plan. Model fit improved from the null model to the final model in the youth ($\Delta DIC_2 = 63.57$, $\Delta WAIC = 62.00$) and caregiver models ($\Delta DIC_2 = 31.99$, $\Delta WAIC = 32.21$).

Unconditional models were estimated to partition the variance in youth and caregiver session engagement across sessions (level 1), clients (level 2), and therapists (level 3). The unconditional models are displayed in Table 26 and 27. Examination of the variance decomposition in youth and caregiver session-level engagement showed notable differences in clustering. Although more of the variance in youth session-level engagement was attributable to the therapist level (23.0%) as opposed to the client level (11.3%), the pattern was flipped for caregiver session-level engagement. Only 7.1% of the variance in caregiver session-level

engagement was attributable to the therapist level, while 22.3% of the variance was attributable to the client level. Substantial session-level variation remained in the youth and caregiver unconditional models respectively (62.2%; 65.7%).

The two full random intercept models for youth and caregiver session-level engagement are displayed in Tables 28-29. The model intercept for both youth and caregiver models indicated an expected session level engagement of 10.41 (95% CrI [9.79, 11.02]) and 12.01 (95% CrI [11.37, 12.66]) respectively at the reference profiles. Hypothesis 3 was supported in both youth and caregiver full models. Total missteps were negatively associated with both youth ($b = -0.12$, 95% CrI [-0.18, -0.06]) and caregiver session-level engagement ($b = -0.10$, 95% CrI [-0.17, -0.03]), controlling for number of youth or caregiver baseline engagement challenges, maximum practice extensiveness, site, and condition. Specifically, each additional misstep occurrence was associated with an expected decrease of 0.12 points in youth session engagement and 0.10 points in caregiver session engagement. Youth baseline engagement challenges, as measured by the number of at-risk items endorsed on the baseline MTT-Y, was not significantly associated with youth session-level engagement. Similarly, caregiver baseline engagement challenges, as measured by the number of at-risk items endorsed in the baseline MTT-CG, was also not significantly associated with caregiver session-level engagement. These findings suggest that the number of engagement challenges the youth and caregiver endorsed at baseline was not reliably associated with their respective aggregate in-session engagement. Site was not significantly associated with either youth or caregiver session-level engagement. Study condition was associated with higher youth session-level engagement for therapists assigned to the PG condition ($b = 0.75$, 95% CrI [0.02, 1.50]); however, the credible interval only marginally excluded zero, suggesting that this finding should be interpreted with caution.

Within the youth model, the addition of predictors and covariates explained additional variance across all three levels, with a notable decrease in unexplained session-level variation to 42.2% from 62.2%. The caregiver model saw similar patterns; however, greater unexplained session-level variation remained (58.9%) following addition of all predictors and covariates. Posterior distributions for credible associations are displayed in Figures 5-9.

Discussion

Building on the prior study, the present study examined how occurrences of therapist missteps may be associated with both the quality of the practices delivered by the therapist in the session (Aim 1), and session-level engagement of the youth and caregiver present in the session (Aim 2).

Treatment Planning

The present study did not find a credible association between therapist missteps and the overall quality of practice delivery for the session. Firstly, the study did not find a significant association between therapist missteps and the average practice extensiveness rating for the session. Additionally, treatment planning (i.e., whether the therapist had identified a treatment plan to deliver in the session) did not appear to be associated with or moderate the association between therapist missteps and the session's average practice extensiveness rating. As such, these findings suggest that the presence or absence of a treatment plan alone does not appear to buffer against or amplify the association between the frequency of missteps and the session's overall quality of practice delivery as operationalized within the present study.

Taken together with findings from Study 1, which found that treatment plan presence and quality was not associated with total misstep occurrences, a notable pattern emerges across studies: simply identifying what to deliver in the next session may not be sufficient to influence

either the occurrence of missteps or the quality of practice delivery. Treatment plans identifying specific practices to deliver as part of a specific client's care have been shown to increase the likelihood that those practices are delivered over the course of treatment (Tsai et al., 2016), but may not ensure greater breadth or depth of delivery. Consistent with these findings, a recent study on trainee implementation of plans established in supervision found that while identifying a treatment plan was associated with higher session adherence, adherence was not associated with therapeutic skill and quality (Maaß et al., 2026). Additional clinical activities emphasizing behavioral rehearsal and experiential learning methods, such as role-playing and modeling, have shown stronger associations with therapist competence and evidence-based treatment delivery (Alfonsson et al., 2020; Choy-Brown et al., 2022; Bjaastad et al., 2025).

Session Engagement

Results from the unconditional models revealed notable differences in how youth and caregiver session-level engagement were structured across therapist, client, and session levels. For youth session engagement, 23.0% of the variance was attributable to therapists, placing this estimate in the upper range of therapist effects reported in recent systematic reviews, which documented therapist effects ranging from 0.2% to 29%, with a weighted average of 5% (Johns et al., 2019). By contrast, only 12.5% of youth session engagement variance was attributable to the client level. This pattern inverted for caregiver session engagement, where 22.3% of the variance was attributable to the client level, compared with only 7.1% at the therapist level. Together, these findings suggest that youth session engagement may be more sensitive to therapist-level factors, whereas caregiver engagement appears to be more sensitive to client-level factors.

Current research is mixed regarding relevant client characteristics that may contribute to the greater client-level variance in caregiver engagement. Although some studies have linked caregiver attendance (a dimension of engagement) to client-level factors such as the youth's presenting problem type, severity, and persistence, age, and gender (Barnett et al., 2020; Wright et al., 2019; Paula et al., 2022); recent systematic reviews utilizing broader definitions of engagement (e.g., caregiver participation engagement, spanning participation in sessions and follow-through on assigned homework) did not find correlations with youth baseline mental health symptoms or demographics (Haine-Schlagel et al., 2022). The greater therapist-level variance in youth engagement is consistent with qualitative findings suggesting that therapist-level factors—including discomfort with client's emotions or approaching difficult topics, training in family therapy, and perceived environmental pressure to produce outcomes quickly—may increase the likelihood of misstep occurrences, while client-level factors such as visible markers of disengagement and more complex needs may also play a role (Park, Becker, Chu et al., 2025).

Across both youth and caregiver session engagement, session-level variance remained substantial in the full models, with the greatest improvement in model fit occurring with the addition of session level variables and minimal improvement with the addition of client and therapist level variables. This is consistent with the field's conceptualization of engagement as dynamic and constructed through “moment-to-moment” interactions, rather than a stable, trait-like characteristic of clients or families. This interpretation is further supported by the absence of credible associations between baseline engagement challenges and session-level engagement for either youth or caregivers, suggesting that what happens within the session itself may be more influential than the engagement challenges clients present initially with. These findings reinforce

the call to action within the field and the rationale for examining within-session predictors—namely, therapist behaviors—as the primary drivers of session-level engagement variability.

Maximum Practice Extensiveness

Analyses examining the maximum practice extensiveness achieved for a session across both aims of the present study converged on a consistent and theoretically meaningful pattern: the depth and breadth with which a therapist delivers an engagement-focused practice within a session has detectable, credible associations with both the occurrence of therapist missteps and the quality of youth and caregiver engagement in that session.

In Aim 1, a credible negative association between the maximum practice extensiveness achieved for the session and total misstep occurrences emerged. Sessions in which therapists delivered at least one practice with high breadth and depth were associated with an expected rate of missteps that was 19% fewer missteps per additional increase in the practice extensiveness rating, after accounting for all other predictors and covariates, including study condition. This finding suggests that higher-quality practice delivery and therapist missteps may inversely co-occur within a session. One interpretation is that therapists who are able to comprehensively deliver a specific practice may be simultaneously less likely to exhibit therapist missteps.

Extending to Aim 2, credible positive associations between maximum practice extensiveness and both youth and caregiver session-level engagement was noted, even after accounting for total misstep occurrences, supervision-to-treatment lag, treatment plan status, site, and study condition. Each one-unit increase in the maximum practice extensiveness for the session was associated with a 0.89 point increase in youth engagement and a 0.63 point increase in caregiver engagement on the 0 to 20 aggregate scale. To contextualize these estimates clinically, we can compare a session in which the most extensively delivered practice received a

rating of 1 (i.e., delivered with minimal depth or breadth) against a session that received a rating of 5 (i.e., nearly all components of a practice delivered). This four-unit difference corresponds to an expected 3.56 point increase in youth engagement and a 2.52 point increase in caregiver engagement. Behaviorally, this point increase within a specific domain corresponds to a shift from low engagement indicators (e.g., repeated minimization of responses, outright refusal to engage with session content) to high engagement indicators (e.g., openly sharing about an experience, demonstrating feeling understood by the therapist).

Taken together, the findings that maximum practice extensiveness was credibly associated across both aims—with fewer missteps in Aim 1, and greater youth and caregiver engagement in Aim 2—suggests that the depth and breadth of practice delivery, measured by the therapist behaviors within the session, may be a meaningful indicator of within-session therapeutic quality more broadly. Conceptually, these findings align with the theoretical framing of therapist misstep as “run-time decisions” made in the absence of clear clinical direction within the session. A therapist who is able to deliver at least one practice with substantial depth and breadth within a session may reflect higher overall preparedness, akin to having a clearer clinical “roadmap” for how to deliver an intervention to target an engagement challenge within the session. This readiness may serve as a protective factor against misstep occurrences more broadly by reducing the likelihood of in-session deviations. Conversely, therapists with lower maximum practice extensiveness may be less equipped to intervene when engagement challenges arise (e.g., client refusal to answer questions), increasing their likelihood of misstep occurrences. However, directionality of the association between the maximum extensiveness rating and misstep occurrences cannot be established from the present analyses. It is possible that therapist preparedness may drive both higher maximum extensiveness rating (i.e., the therapist’s

ability to deliver a practice with high quality) and lower misstep occurrences when navigating engagement challenges within the session. Future analyses examining the temporal unfolding of therapist missteps and the occurrence and quality of evidence-based strategies delivered by the therapist may allow for a more granular understanding of the observed association.

Therapist Missteps

As hypothesized, total misstep count was negatively associated with both youth and caregiver session engagement. Each additional misstep corresponded to an expected -0.12 point change for youth and a -0.10 change for caregiver engagement. The small per-misstep effect is theoretically expected, as a single misstep is not anticipated to substantially alter engagement. Consistent with prior research, these findings suggest that occurrence of therapist missteps, even when infrequent, may have a detectable and credible association with how clients and caregivers engaged with treatment in that session. Each misstep occurrence is associated with a modest but consistent downward shift, such that youth and caregivers may become incrementally less responsive and participatory over the course of a session. Prior studies have found that even low levels of negative therapist behaviors, such as making critical, judgmental, countering, or minimizing comments have been significantly associated with early termination from services prior to completing four sessions, suggesting that clients may respond to such missteps by disengaging from treatment over time (Elkin et al., 2014). This is consistent with community partners' predictions that youth and families may respond to missteps by "just dropping out" of treatment (Park, Becker, Chu et al., 2025). Overall, these findings reinforce that less optimal therapist behaviors, even at relatively low levels, may be sufficient to impede clinical progress (Binder & Strupp, 1997). This may be especially consequential when missteps are systematic and repeated without adequate repair by the therapist (Rutan, 2021). This is consistent with Luo

and Levendosky's (2025) "every moment matters" hypothesis, which proposes that the accumulation of momentary, within-session interactions between the therapist and client contributes meaningfully to session-level outcomes, and by extension, broader treatment outcomes over time.

Study Condition

Study condition remained credibly associated with total misstep occurrences, even with the addition of maximum practice extensiveness to the model. Therapists in the Practice Guidelines (PG) condition were associated with an expected rate of misstep occurrences that was approximately three times higher than the expected rate for therapists in the Coordinated Knowledge System (CKS) condition, accounting for treatment plan status, supervision-to-treatment lag, maximum practice extensiveness, and site. Additionally, 27.4% of variance in therapist missteps remained at the therapist level in the full model. Together, these findings suggest that the provision of coordinated decision-making supports constrained the occurrence of missteps above and beyond supporting therapists in delivering at least one practice with high depth and breadth.

The magnitude of the study condition effect on total misstep occurrences is notable, given that both conditions were provided information about engagement concerns and knowledge of the same 11 evidence-based engagement practices. What distinguished the CKS condition was not knowledge of what practices existed or whether engagement challenges were present, but structured support for how to reason about and act on that knowledge, including a framework embedded within supervision for detecting and prioritizing engagement concerns, guidance on matching specific practices to specific problems, brief, step-by-step instructions on how to deliver the 11 engagement practices, and ongoing monitoring of practice delivery (Chorpita et

al., 2025). In the absence of supports to identify what specific engagement challenges were reported, how to match practices to problems, and actionable guidance on how to deliver the practices, therapists may have faced greater uncertainty about how to intervene—conditions under which the likelihood of missteps, conceptualized as in-the-moment deviations from evidence-based responses, may be elevated.

In Aim 2, study condition was not credibly associated with caregiver session engagement, and credibly associated with youth session engagement; however the credible interval marginally excluded zero. One possibility is that the aggregate engagement scores may not have been sensitive enough to detect targeted improvements in specific engagement domains. Therapists in the CKS condition were provided with tools and supports to select specific domains of engagement to target in their therapy sessions, as opposed to attempting to intervene on all domains of engagement that were flagged as at risk. As such, effects of study condition may have been attenuated when session engagement was aggregated across the REACH domains, and may have been better captured by examining domain-specific session engagement. Additionally, the pathway through which structured clinical decision-making supports may impact session-level engagement is likely through the quality of practice delivery in the session. Future work examining domain-specific engagement outcomes may better capture the effect of structured clinical decision-making supports on how youth and caregiver engagement may manifest within a session.

Limitations

This study is among few studies to examine how therapist missteps relate to session-level outcomes in youth mental health treatment settings. Despite this strength, several limitations warrant consideration. The complex nested structure of the data and the small sample size at

higher levels may have limited statistical power and the number of predictors that could be included in final models, reflected in the substantial unexplained session-level variance that remained. Additionally, the correlational design limited the ability to rule out potential unmeasured variables associated with both predictor and outcomes that may account for some observed associations. The present analyses also did not examine how total misstep occurrences may differentially impact specific domains of engagement, or how specific types of missteps (e.g., advice-giving, arguing) may impact both global versus specific domains of engagement. Community partners have noted that certain “missteps,” such as the use of humor, may be effective in certain contexts (e.g., in the presence of strong therapeutic alliance) and potentially invalidating and dismissive in others (Park, Becker, Chu et al., 2025). Given the limited prior research, an aggregate engagement measure was utilized to establish whether a credible general association between misstep occurrences and session engagement could be detected. Future research could examine how misstep occurrences may be differentially related to engagement and other treatment or process outcomes alongside other notable therapist behaviors in the session, such as use of specific strategies measured in Study 1 (e.g., expressing empathy).

The use of average practice extensiveness as the Aim 1 primary outcome similarly limited interpretability and ability to detect effects by masking meaningful within-session variation in treatment delivery relevant to the study aims. Two sessions can yield similar average extensiveness ratings despite representing clinically distinct patterns of evidence-based practice delivery. For example, a session in which the therapist delivers a single practice with substantial depth (e.g., providing psychoeducation about depression in great depth and detail) alongside a brief check-in on barriers to attendance may be scored similarly as a session in which the therapist delivered several practices with only moderate breadth (e.g., jumping between limited

psychoeducation about services, preliminary goal-setting, and a brief check-in about attendance). This conflation limits the interpretive precision of the average extensiveness rating as an indicator of treatment quality and may partly account for the unexpected null finding that study condition was not significantly associated with the average extensiveness rating—despite the intervention’s explicit focus on guiding therapists in the selection and delivery of evidence-based practices in the session. In contrast, the maximum extensiveness rating appeared to be a more sensitive and clinically meaningful indicator of practice delivery quality, consistent with findings by Lau and colleagues (2024), who found that the depth of delivery of the therapist’s most extensiveness used practice, even when averaged across sessions, was predictive of treatment outcomes. Future work could consider using similarly granular metrics of therapist behaviors related to practice delivery to capture how the quality of treatment delivered may present and fluctuate within sessions.

Conclusion

Across the two studies, this research provides an integrated examination of therapist missteps in youth publicly funded mental health settings, examining how and when they occur, and their associations with treatment delivery quality and engagement. Study 1 established that inadvertent therapist missteps occur routinely in youth mental health services and can be measured reliably in observational data. Additionally, Study 1 established that provision of coordinated clinical decision-making supports were associated with substantially lower expected rates of missteps. Study 2 extended these findings by demonstrating that the misstep occurrences were credibly associated with expected decreases in both overall practice quality and youth and caregiver engagement in the session. Variance decomposition further revealed that a meaningful proportion of variance in youth session engagement was attributable at the therapist level and

session level, positioning in-session therapist behaviors as a clinically meaningful target for future research. Together, these findings reinforce a conceptualization of therapist missteps as predictable features of treatment delivery that appear to be responsive to clinical decision-making support. Overall, these studies highlight the value of equipping therapists with clinical decision-making supports that can ensure predictable blind curves are well-illuminated and guardrails are structurally embedded, so that missteps are less likely to occur and therapists are positioned to respond with precision.

Tables

Table 1

Misstep Codes and Definitions

Misstep Code	Definition
Advice-giving	Giving advice or uses direct persuasion by telling the recipient what to do in a given situation, but without teaching, reinforcing, or consolidating generalizable skill.
Arguing or displaying doubt/disbelief	Arguing or expressing diverging views, typically in a heated or persistent manner (e.g., persistent non-Socratic questioning, fact-checking).
Criticizing	Indicating perceived faults of recipient in a disapproving and judgmental way (e.g., preemptively blaming/accusing, negatively labeling words or behaviors, pointing out mistakes non-constructively, complaining about recipient).
Demanding attention/engagement	Engaging in verbal bids for attention or engagement from the recipient (e.g., interrupting, talking over, commanding, commenting on lack of engagement).
Downplaying problem, identity, beliefs	Minimizing or trivializing problem, identity, or beliefs (e.g., putting a “positive spin” on a problem, oversimplifying the problem and barriers to solutions).
Eliciting shame/guilt	Using feelings of guilt, remorse, or a sense of responsibility about recipients’ words/behaviors to change how they think, feel, or might act (e.g., implying recipient “owes” change to others, expressing disappointment, highlighting negative impact/burden of recipient’s words/behaviors on others).
Highlighting negative consequences/ appealing to fear to motivate change	Highlighting negative consequences or attempting to increase fear to motivate behavior change in the recipient.
Inappropriate elicitation/disclosure of information	Eliciting or disclosing information that is irrelevant to the family and their treatment (e.g., eliciting or disclosing personal information or anecdotes in a lengthy manner without tying back to client concerns, gossiping).
Making a joke at family’s expense	Using humor directed at the recipient or recipient’s situation using inappropriate humor (e.g., making recipient/situation the target of the joke, reacting inappropriately, commenting humorously).
Speculating or making assumptions	Assuming or speculating on recipient’s presenting problems without evidence/information or checking in with the recipient.

Note. Adapted from Park, Becker, Chu et al., 2025.

Table 2*Frequency and Mean (SD) of Missteps Observed within Sessions by Participant Type*

Missteps Observed	Total (% of 979 occurrences)	Session Participant Type					
		Youth Only (231 events)		Caregiver Only (35 events)		Family (125 events)	
		<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>
Advice-giving	369 (37.9%)	189	0.82 (1.6)	31	0.89 (1.6)	149	1.19 (2.0)
Arguing or displaying doubt/disbelief	152 (15.5%)	75	0.32 (0.9)	7	0.20 (0.5)	70	0.56 (1.2)
Criticizing	109 (11.1%)	44	0.19 (0.5)	5	0.14 (0.5)	60	0.48 (1.7)
Demanding attention	44 (4.5%)	22	0.10 (0.4)	0	0	22	0.18 (0.8)
Downplaying problem, identity, beliefs	56 (5.7%)	42	0.18 (0.7)	1	0.3 (0.2)	13	0.10 (0.4)
Eliciting shame/guilt	34 (3.5%)	14	0.06 (0.4)	0	0	20	0.16 (0.8)
Highlighting negative consequences/ appealing to fear to motivate change	71 (7.3%)	40	0.17 (0.5)	2	0.06 (0.2)	29	0.23 (0.8)
Inappropriate elicitation/disclosure of information	52 (5.3%)	22	0.10 (0.7)	2	0.06 (0.2)	28	0.22 (0.6)
Making a joke at family's expense	36 (3.7%)	10	0.04 (0.2)	0	0	26	0.21 (0.5)
Speculating or making assumptions	56 (5.7%)	44	0.19 (0.6)	1	0.03 (0.2)	11	0.09 (0.4)
Any Misstep	979	502	2.17 (3.6)	49	1.4 (2.4)	428	3.42 (5.9)

Note. Adapted from Park, Becker, Chu et al., 2025.

Table 3*Frequency of Missteps in Family Sessions by Recipient*

Missteps Observed	Youth-Directed	Caregiver-Directed	Family-Directed
Advice-giving	84 (56%)	62 (42%)	3 (2%)
Arguing or displaying doubt/disbelief	53 (76%)	16 (23%)	1 (1%)
Criticizing	46 (77%)	10 (17%)	4 (7%)
Demanding attention	18 (82%)	4 (18%)	0
Downplaying problem, identity, beliefs	10 (77%)	2 (15%)	1 (8%)
Eliciting shame/guilt	18 (90%)	2 (10%)	0
Highlighting negative consequences/ appealing to fear to motivate change	25 (86%)	4 (14%)	0
Inappropriate elicitation/disclosure of information	10 (36%)	16 (57%)	2 (7%)
Making a joke at family's expense	17 (65%)	3 (12%)	6 (23%)
Speculating or making assumptions	3 (27%)	5 (45%)	3 (27%)
Any Misstep	284 (66%)	124 (29%)	20 (5%)

Note. Adapted from Park, Becker, Chu et al., 2025. Percentages are calculated using the total number of misstep occurrences for each misstep type within family sessions (n = 125 sessions; 428 total misstep occurrences).

Table 4*Frequency of Therapist Missteps by Immediately Preceding Therapist Behavior Type.*

Missteps Observed	Preceding Therapist Behavior		
	Clinical Procedure (% of 632 instances)	Any Misstep (% of 104 instances)	No Codes Applicable (% of 269 instances)
Advice-giving	254 (40.6%)	33 (31.7%)	82 (30.5%)
Arguing or displaying doubt/disbelief	103 (16.5%)	12 (11.5%)	37 (13.8%)
Criticizing	66 (10.5%)	13 (12.5%)	30 (11.2%)
Demanding attention	32 (5.1%)	3 (2.9%)	9 (3.3%)
Downplaying problem, identity, beliefs	34 (5.4%)	7 (6.7%)	15 (5.6%)
Eliciting shame/guilt	16 (2.6%)	5 (4.8%)	13 (4.8%)
Highlighting negative consequences/ appealing to fear to motivate change	38 (6.1%)	16 (15.4%)	17 (6.3%)
Inappropriate elicitation/disclosure of information	27 (4.3%)	5 (4.8%)	20 (7.4%)
Making a joke at family's expense	26 (4.2%)	2 (1.9%)	28 (10.4%)
Speculating or making assumptions	30 (4.8%)	8 (7.7%)	18 (6.7%)
Any Misstep (% of 979 instances)	632 (63.2%)	104 (10.6%)	269 (27.5%)

Note. Adapted from Park, Becker, Chu et al., 2025.

Table 5*Descriptive Statistics for Total Missteps Model Variables*

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max
Total Missteps	391	2.51	4.44	0	42
Supervision-to-Treatment Lag	391	9.00	10.57	0	70
Treatment Plan Status					
No Plan	123 (31.5%)	—	—	—	—
Non-Engagement Plan	82 (21.0%)	—	—	—	—
Unmatched Engagement Plan	72 (18.4%)	—	—	—	—
Matched Plan	114 (29.2%)	—	—	—	—
Site (% California)	213 (45.5%)	—	—	—	—
Study Condition (% CKS)	210 (53.7%)	—	—	—	—

Table 6*Overdispersion Statistics for Total Missteps Variable*

Statistic	Estimate	Interpretation
Variance	19.63	
Variance to Mean Ratio	7.82	Overdispersed
Skewness	3.81	Right-skewed

Table 7*Distribution of Total Missteps Across Levels of Treatment Plan Status*

Treatment Plan Status	<i>N</i>	Mean Missteps	Median Missteps	% No Missteps
No Plan	123	2.04	1	43.1%
Non-Engagement Plan	72	3.15	2	38.9%
Unmatched Plan	82	2.78	1	42.7%
Matched Plan	114	2.42	1	44.7%

Table 8*Correlations For Continuous Session-Level Total Missteps Model Variables*

Variable	1	2	3
Total Missteps	--		
Supervision-to-Treatment Lag	.17* [.07, .26]	--	
Maximum Practice Extensiveness	-.11* [-.21, -.02]	-.12* [-.22, -.02]	--

Table 9*Convergence Diagnostics and Model Fit Statistics for Total Missteps Random Intercept Models**(Study 1 Aim 2)*

Model	Burn-In Iterations	Final Iterations	Highest PSRF	Lowest Effective <i>N</i>	DIC ₂	WAIC
Null Model	20,000	30,000	1.02	127.15	-237.74	-235.87
Level 1 Variables Added	30,000	30,000	1.04	138.41	-237.95	-236.04
Level 3 Variables Added	30,000	30,000	1.02	138.52	-236.97	-235.43

Note. PSRF = Potential Scale Reduction Factor, DIC₂ = Deviance Information Criterion specified

for multilevel models with missing data, WAIC = Watanabe-Akaike Information Criterion. The highest PSRF reported corresponds to the highest PSRF at the final burn-in checkpoint.

Level 1 variables added: treatment plan status (no plan, non-engagement plan, unmatched engagement plan with matched engagement plan as reference group); supervision-to-treatment lag (grand-mean centered)

Level 3 variables added: site, study condition

Table 10

Unconditional three-level negative binomial model of therapist missteps (Study 1 Aim 2)

Parameter	Estimate	SE	95% CrI	Latent ICC
Intercept (IRR)	0.84	0.34	[0.40, 1.72]	
Variance Components				
Therapist-Level	1.43	0.38	[0.86, 2.34]	29.7%
Client-Level	0.09	0.10	[0.01, 0.36]	1.8%
Distribution Parameter	0.86	0.14	[0.61, 1.17]	

Note. IRR = Incidence rate ratio. CrI = Credible interval. ICC = Intraclass Correlation

Coefficients. The latent variable approach for intraclass correlation (ICC) calculation in count models was utilized to derive ICCs (Hox et al., 2017).

Table 11*Full Random Intercept Model for Lag, Plan Status, and Cross-Level Interactions (Study 1 Aim 2)*

Parameter	Estimate	95% CrI	Latent ICC
Coefficients			
Intercept	0.52*	[0.28, 0.88]	
Unmatched Engagement Plan	1.76	[0.90, 3.39]	
Non-Engagement Plan	1.20	[0.57, 2.52]	
No Plan	1.32	[0.71, 2.43]	
Lag (days)	1.01	[1.00, 1.02]	
Site	1.79*	[1.02, 3.13]	
Study Condition	3.18*	[1.46, 6.82]	
Unmatched Engagement Plan*Control Condition	0.51	[0.21, 1.24]	
Non-Engagement Plan* Control Condition	0.87	[0.34, 2.24]	
No Plan*Control Condition	0.72	[0.32, 1.65]	
Variance Components			
Therapist-Level	1.10	[0.62, 1.88]	24.5%
Client-Level	0.10	[0.01, 0.36]	2.2%
Distribution Parameter	0.89	[0.62, 1.23]	

Note. Estimates for exponentiated coefficients are Incidence Rate Ratios (IRRs). * = 95%

credible interval for exponentiated coefficients excluded 1.0. The latent variable approach for intraclass correlation (ICC) calculation in count models was utilized to derive ICCs (Hox et al., 2017).

Table 12*Descriptive Statistics for Average Practice Extensiveness Model Variables*

Variable	<i>n</i>	<i>M</i>	<i>SD</i>	Max	Min
Average Extensiveness	391	1.72	0.65	0	5
Total Missteps	391	2.53	4.44	0	42
Total Practices	391	4.23	2.07	0	11
Plan Made (% yes)	391 (68.5%)	–	–	–	–
Site (% California)	213 (45.5%)	–	–	–	–
Study Condition (% CKS)	210 (53.7%)	–	–	–	–

Table 13*Distribution of Average Extensiveness Ratings Across Site, Study Condition, and Plan Made*

Variable	<i>N</i>	<i>M</i>	<i>SD</i>
Site			
California (0)	214	1.83	0.63
South Carolina (1)	177	1.59	0.65
Study Condition			
CKS (0)	209	1.74	0.58
PG (1)	182	1.70	0.72
Plan Made			
Plan Made (0)	267	1.76	0.61
No Plan (1)	124	1.64	0.72

Table 14*Correlations for Average Practice Extensiveness Model Variables*

Variable	1	2	3
Average Extensiveness	–		
Total Missteps	-.13* [-.22, -.03]	–	
Total Practices	.16* [.07, .26]	.04 [-.06, .14]	–

Table 15*Convergence Diagnostics and Model Fit Statistics for Average Extensiveness Random Intercept**Models (Study 2 Aim 1)*

Model	Burn-In Iterations	Final Iterations	Highest PSRF	Lowest Effective <i>N</i>	Marginal DIC ₂	Marginal WAIC
Null	20,000	20,000	1.02	159.16	765.54	765.75
Level 1 Variables Added	20,000	20,000	1.03	153.39	751.72	752.68
Level 3 Variables Added	30,000	30,000	1.01	136.30	752.93	754.67

Note. PSRF = Potential Scale Reduction Factor, DIC₂ = Deviance Information Criterion specified

for multilevel models with missing data, WAIC = Watanabe-Akaike Information Criterion. The highest PSRF reported corresponds to the highest PSRF at the final burn-in checkpoint.

Level 1 variables added: Total misstep occurrences, total practices delivered in the session (grand mean centered), plan status (in binary form; 0 = presence of treatment plan)

Level 3 variables added: site, study condition

Table 16*Unconditional Three-Level Model of Average Practice Extensiveness*

Parameter	Estimate	95% CrI
Intercept	1.73*	[1.51 1.94]
Variance Components		
Therapist-Level	0.07	[0.020, 0.134]
Client-Level	0.04	[0.005, 0.111]
Proportion Variance Explained (R^2)		
By Client-Level Random Intercepts	9.4%	[1.1%, 24.6%]
By Therapist-Level Random Intercepts	15.0%	[4.7%, 27.9%]
By Session-Level Residual Variation	73.8%	[60.1%, 85.6%]

Note. CrI = Credible interval. * = 95% credible interval excluded 0.

Table 17*Full Random Intercept Model for Average Practice Extensiveness (Study 2 Aim 1)*

Parameter	Estimate (SE)	95% CrI
Coefficients		
Intercept	1.87* (0.07)	[1.73, 2.01]
Total Missteps	-0.02* (0.01)	[-0.04, -0.01]
Total Practices	0.02 (0.02)	[-0.01, 0.05]
Site	-0.16 (0.09)	[-0.34, 0.01]
Study Condition	0.03 (0.09)	[-0.14, 0.20]
Plan Made	-0.09 (0.08)	[-0.25, 0.08]
Total Missteps x Plan Made	0.004 (0.02)	[-0.03, 0.04]
Variance Components		
Therapist-Level	0.06	[0.015, 0.124]
Client-Level	0.03	[0.001, 0.090]
Proportion Variance Explained (R^2)		
By Fixed Effects	6.6%	[2.5%, 12.5%]
By Client-Level Random Intercepts	6.9%	[0.3%, 19.7%]
By Therapist-Level Random Intercepts	13.4%	[3.4%, 25.4%]
By Session-Level Residual Variation	71.8%	[59.2%, 83.8%]

Note. CrI = Credible interval. * = 95% credible interval excluded 0.

Table 18*Convergence Diagnostics and Model Fit Statistics for Total Missteps Random Intercept Models**(Study 2 Aim 1 Exploratory)*

Model	Burn-In Iterations	Final Iterations	Highest PSRF	Lowest Effective <i>N</i>	DIC ₂	WAIC
Null	20,000	30,000	1.02	127.15	-237.74	-235.87
Level 1 Variables Added	40,000	40,000	1.01	164.98	-243.72	-241.73
Level 3 Variables Added	40,000	50,000	1.01	137.27	-242.17	-240.47

Note. PSRF = Potential Scale Reduction Factor, DIC₂ = Deviance Information Criterion specified

for multilevel models with missing data, WAIC = Watanabe-Akaike Information Criterion. The highest PSRF reported corresponds to the highest PSRF at the final burn-in checkpoint.

Level 1 variables added: treatment plan status (no plan, non-engagement plan, unmatched engagement plan with matched engagement plan as reference group); supervision-to-treatment lag (grand mean centered), maximum practice extensiveness (grand mean centered)

Level 3 variables added: site, study condition

Table 19

Full Random Intercept Model for Lag, Plan Status, and Cross-Level Interactions (Study 2 Aim 1 Exploratory)

Parameter	Estimate	95% CrI	Latent ICC
Coefficients			
Intercept	0.54	[0.28, 1.02]	
Unmatched Engagement Plan	1.74	[0.90, 3.36]	
Non-Engagement Plan	1.21	[0.58, 2.57]	
No Plan	1.35	[0.74, 2.46]	
Lag (days)	1.01	[0.99, 1.02]	
Maximum Practice Extensiveness	0.81*	[0.70, 0.96]	
Site	1.57	[0.86, 2.87]	
Study Condition	3.01*	[1.39, 6.65]	
Unmatched Engagement Plan*PG	0.53	[0.22, 1.26]	
Non-Engagement Plan*PG	0.85	[0.34, 2.16]	
No Plan*PG	0.73	[0.33, 1.63]	
Variance Components			
Therapist-Level	1.28	[0.736, 2.154]	27.4%
Client-Level	0.10	[0.002, 0.382]	2.2%
Distribution Parameter	0.79	[0.533, 1.123]	

Note. Estimates for exponentiated coefficients are Incidence Rate Ratios (IRRs). CrI = Credible interval. * = 95% credible interval for exponentiated coefficients excluded 1.0. The latent variable approach for intraclass correlation (ICC) calculation in count models was utilized to derive ICCs (Hox et al., 2017). PG = Practice Guidelines.

Table 20*Descriptive Statistics For Youth Engagement Model Variables*

Variable	<i>n</i>	<i>M</i>	<i>SD</i>	Min	Max
Youth Session Engagement	356	10.16	2.80	1	19
Total Missteps	356	2.64	4.58	0	42
Maximum Practice Extensiveness	356	2.60	1.16	0	5
Number of At-Risk Baseline MTT Items	339	4.89	4.92	1	20
Site (% SC)	356 (46.1%)	—	—	—	—
Study Condition (% PG)	356 (47.2%)	—	—	—	—

Table 21*Descriptive Statistics For Caregiver Engagement Model Variables*

Variable	<i>n</i>	<i>M</i>	<i>SD</i>	Min	Max
Caregiver Session Engagement	160	2.45	2.45	3	18
Total Missteps	160	2.99	5.38	0	42
Maximum Practice Extensiveness	160	2.48	1.20	0	5
Number of At-Risk Baseline MTT Items	156	2.94	5.67	1	14
Site (% SC)	40.61%	—	—	—	—
Study Condition (% PG)	46.9%	—	—	—	—

Table 22*Correlations For Session-Level Youth Engagement Model Variables*

Variable	1	2	3
Youth Session Engagement	–		
Total Missteps	-.22* [-.32, -.12]	–	
Maximum Practice Extensiveness	.43* [.34, .51]	-.14* [-.24, -.04]	–

Table 23*Correlations For Session-Level Caregiver Engagement Model Variables*

Variable	1	2	3
Caregiver Session Engagement	–		
Total Missteps	-.28* [-.42, -.13]	–	
Maximum Practice Extensiveness	.43* [.19, .48]	-.04 [-.20, .12]	–

Table 24*Convergence Diagnostics and Model Fit Statistics for Youth Engagement Random Intercept**Models (Study 2 Aim 2)*

Model	Burn-In Iterations	Final Iterations	Highest PSRF	Lowest Effective <i>N</i>	DIC₂	WAIC
Null Model	10,000	20,000	1.01	124.76	1716.17	1716.76
Level 1 Variables Added	10,000	20,000	1.01	286.12	1652.36	1653.49
Level 2 Variables Added	10,000	20,000	1.03	436.73	1640.84	1642.18
Level 3 Variables Added	20,000	20,000	1.03	208.20	1652.60	1654.76

Note. PSRF = Potential Scale Reduction Factor, DIC₂ = Deviance Information Criterion specified for multilevel models with missing data, WAIC = Watanabe-Akaike Information Criterion. The highest PSRF reported corresponds to the highest PSRF at the final burn-in checkpoint.

Level 1 variables added: total misstep occurrences, maximum practice extensiveness (grand mean centered)

Level 2 variables added: number of at-risk MTT-Y items endorsed at baseline (grand mean centered)

Level 3 variables added: site, study condition

Table 25

Convergence Diagnostics and Model Fit Statistics for Caregiver Engagement Random Intercept Models (Study 2 Aim 2)

Model	Burn-In Iterations	Final Iterations	Highest PSRF	Lowest Effective <i>N</i>	DIC₂	WAIC
Null Model	10,000	20,000	1.02	141.59	759.62	766.14
Level 1 Variables Added	10,000	20,000	1.02	118.96	728.48	734.36
Level 2 Variables Added	20,000	20,000	1.02	195.84	729.32	735.89
Level 3 Variables Added	80,000	20,000	1.04	118.52	727.63	733.93

Note. PSRF = Potential Scale Reduction Factor, DIC₂ = Deviance Information Criterion specified for multilevel models with missing data, WAIC = Watanabe-Akaike Information Criterion. The highest PSRF reported corresponds to the highest PSRF at the final burn-in checkpoint.

Level 1 variables added: total misstep occurrences, maximum practice extensiveness (grand mean centered)

Level 2 variables added: number of at-risk MTT-CG items endorsed at baseline (grand mean centered)

Level 3 variables added: site, study condition

Table 26*Unconditional Three-Level Model of Youth Session-Level Engagement*

Parameter	Estimate	SE	95% CrI
Intercept	10.90*	0.51	[9.88, 11.90]
Variance Components			
Therapist-Level	1.89	0.66	[0.74, 3.35]
Client-Level	0.93	0.61	[0.08, 2.34]
Proportion Variance Explained (R^2)			
By Client-Level Random Intercepts	11.3%		[0.9%, 27.6%]
By Therapist-Level Random Intercepts	23.0%		[9.5%, 36.3%]
By Session-Level Residual Variation	62.2%		[48.6%, 76.6%]

Note. CrI = Credible interval. * = 95% credible interval for exponentiated coefficients excluded

0.

Table 27*Unconditional Three-Level Model of Caregiver Session-Level Engagement*

Parameter	Estimate	SE	95% CrI
Intercept	12.37*	0.56	[11.25, 13.45]
Variance Components			
Therapist-Level	0.46		[0.02, 1.64]
Client-Level	1.45		[0.12, 3.14]
Proportion Variance Explained (R^2)			
By Client-Level Random Intercepts	22.3%		[1.9%, 42.8%]
By Therapist-Level Random Intercepts	7.1%		[0.3%, 22.9%]
By Session-Level Residual Variation	65.7%		[47.3%, 87.9%]

Note. CrI = Credible interval. * = 95% credible interval for exponentiated coefficients excluded

0.

Table 28*Final Random Intercept Model for Youth Session-Level Engagement*

Parameter	Estimate	SE	95% CrI
Coefficients			
Intercept	10.41*	0.32	[9.79, 11.02]
Total Missteps	-0.12*	0.03	[-0.18, -0.06]
Max Practice Extensiveness	0.89*	0.12	[0.66, 1.12]
Number of At-Risk Baseline MTT Items	-0.02	0.03	[-0.08, 0.05]
Site	-0.65	0.40	[-1.43, 0.10]
Study Condition	0.75*	0.38	[0.02, 1.50]
Variance Components			
Therapist-Level	1.25	0.52	[0.44, 2.42]
Client-Level	1.46	0.59	[0.43, 2.67]
Proportion Variance Explained (R^2)			
By Fixed Effects	24.3%		[16.7%, 32.2%]
By Client-Level Random Intercepts	17.3%		[5.1%, 31.1%]
By Therapist-Level Random Intercepts	14.7%		[4.4%, 26.5%]
By Session-Level Residual Variation	42.2%		[32.7%, 55.4%]

Note. CrI = Credible interval. * = 95% credible interval for exponentiated coefficients excluded 0.

Table 29*Final Random Intercept Model for Caregiver Session-Level Engagement*

Parameter	Estimate	SE	95% CrI
Coefficients			
Intercept	12.01*	0.33	[11.37, 12.66]
Total Missteps	-0.10*	0.04	[-0.17, -0.03]
Max Practice Extensiveness	0.63*	0.16	[0.33, 0.94]
Number of At-Risk Baseline MTT Items	0.03	0.04	[-0.04, 0.10]
Site	-0.77	0.43	[-1.60, -0.09]
Study Condition	0.56	0.41	[-0.27, 1.36]
Variance Components			
Therapist-Level	0.25	0.30	[0.009, 1.103]
Client-Level	0.91	0.58	[0.037, 2.234]
Proportion Variance Explained (R^2)			
By Fixed Effects	21.3%		[11.2%, 32.3%]
By Client-Level Random Intercepts	13.9%		[0.6%, 31.7%]
By Therapist-Level Random Intercepts	3.9%		[0.1%, 15.7%]
By Session-Level Residual Variation	58.9%		[42.8%, 76.6%]

Note. CrI = Credible interval. * = 95% credible interval for exponentiated coefficients excluded 0.

Figures

Figure 1

Posterior Distribution of the Effect of Site on Total Missteps (Study 1 Aim 2)

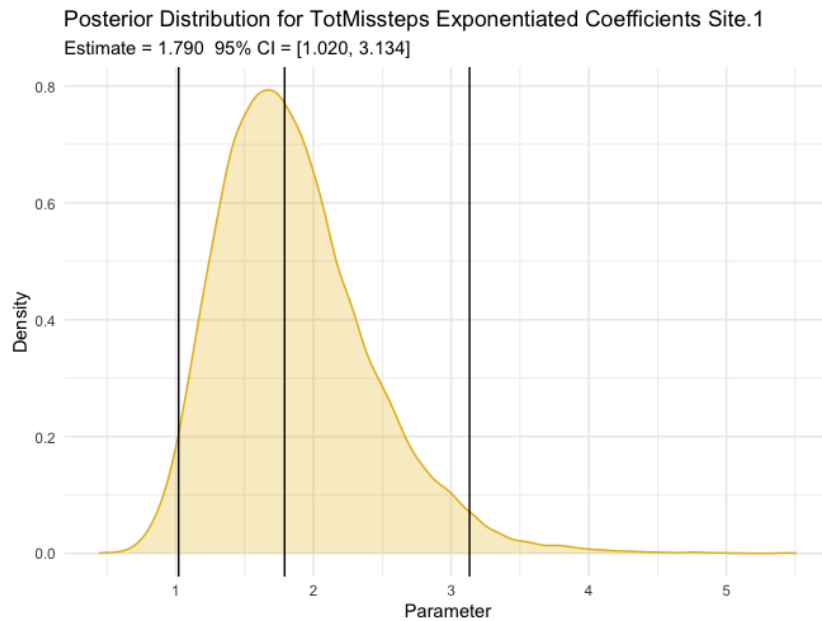


Figure 2

Posterior Distribution of the Effect of Study Condition on Total Missteps (Study 1 Aim 2)

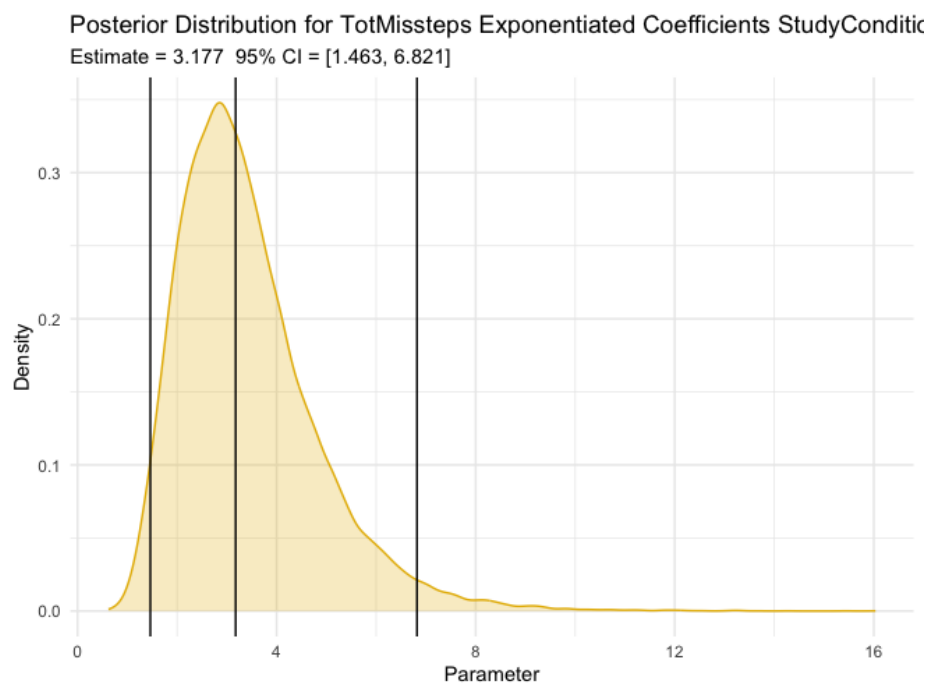


Figure 3

Posterior Distribution of the Effect of Maximum Practice Extensiveness on Total Missteps (Exploratory)

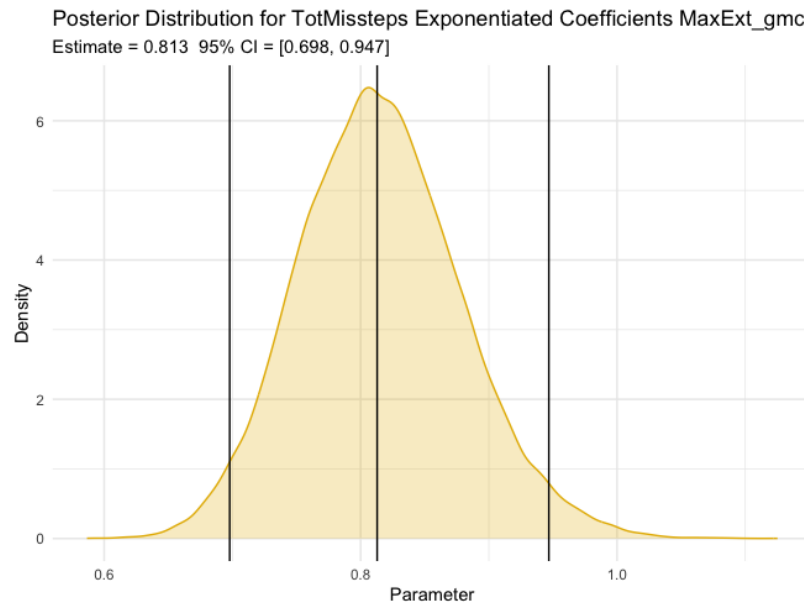


Figure 4

Posterior Distribution of the Effect of Study Condition on Total Missteps (Exploratory)

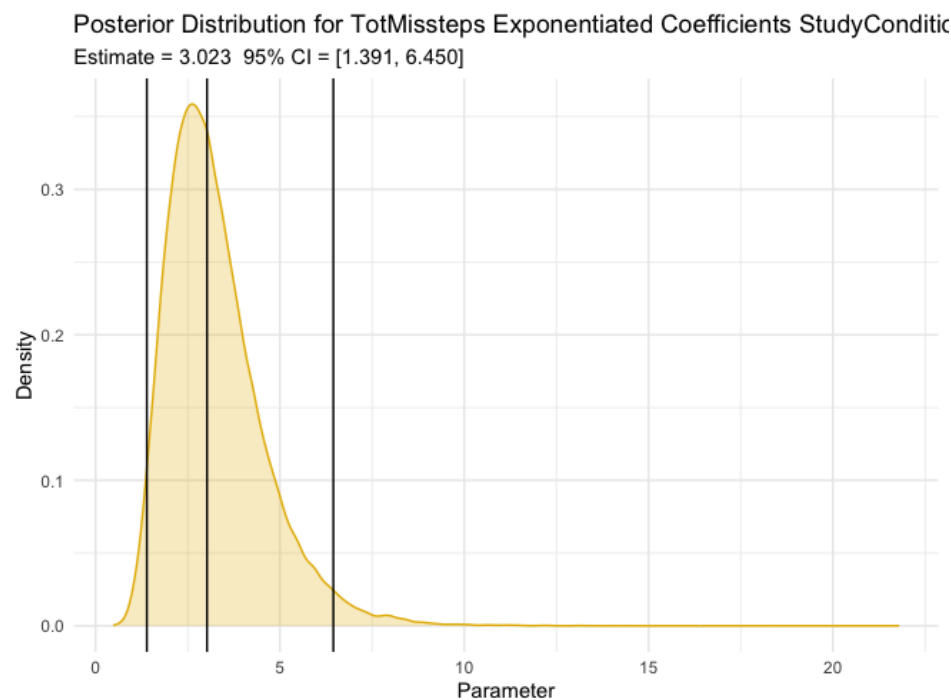


Figure 5

Posterior Distribution of the Effect of Total Missteps on Youth Session-Level Engagement (Study 2 Aim 2)

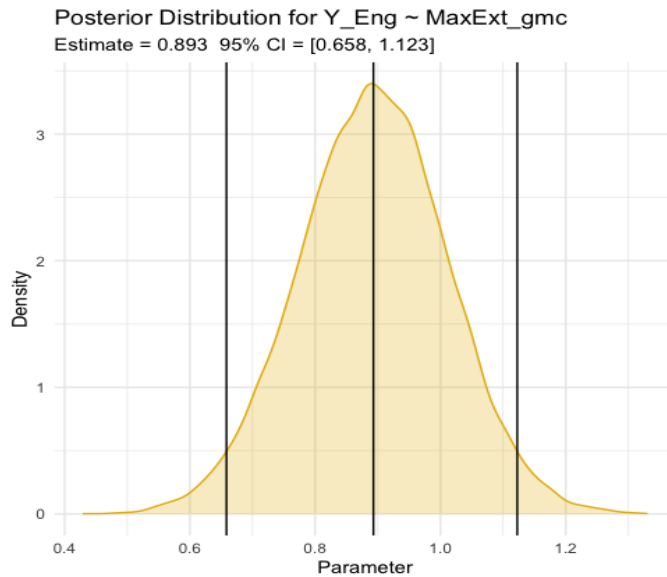


Figure 6

Posterior Distribution of the Effect of Maximum Practice Extensiveness on Youth Session-Level Engagement (Study 2 Aim 2)

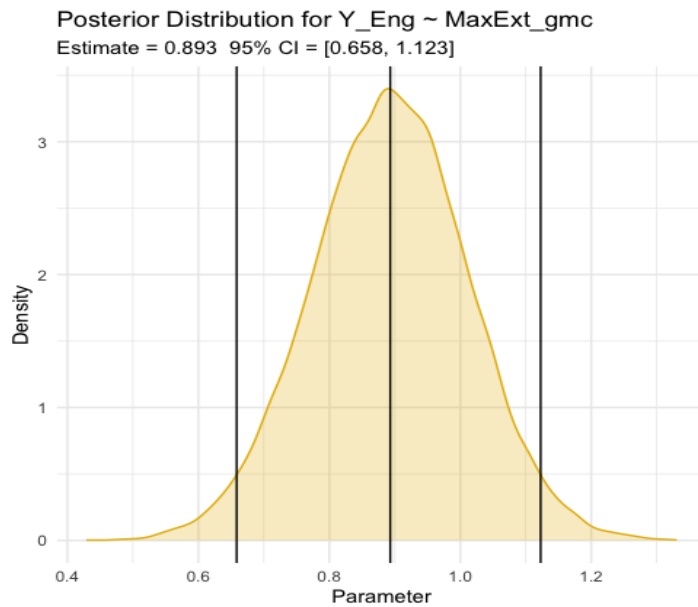


Figure 7

Posterior Distribution of the Effect of Study Condition on Youth Session-Level Engagement

(Study 2 Aim 2)

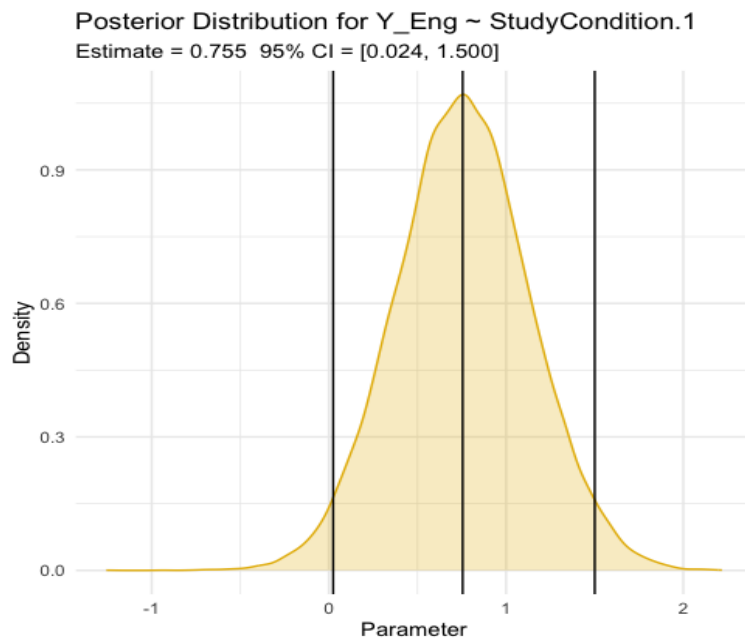


Figure 8

Posterior Distribution of the Effect of Total Missteps on Caregiver Session-Level Engagement

(Study 2 Aim 2)

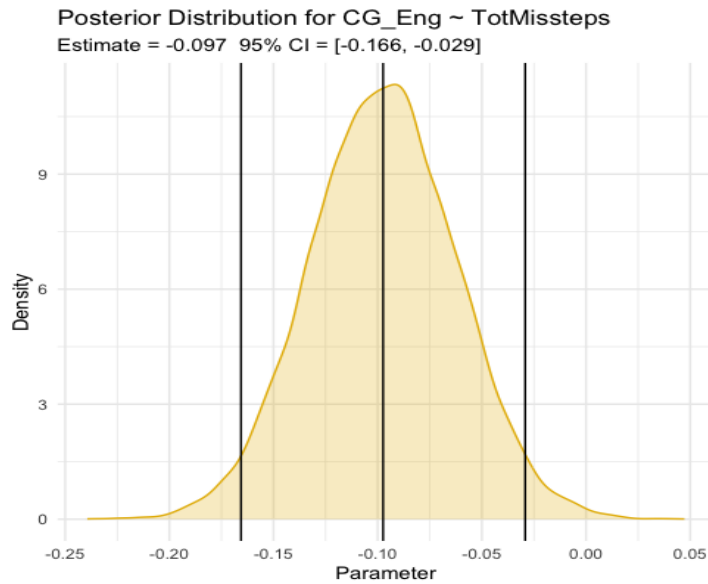
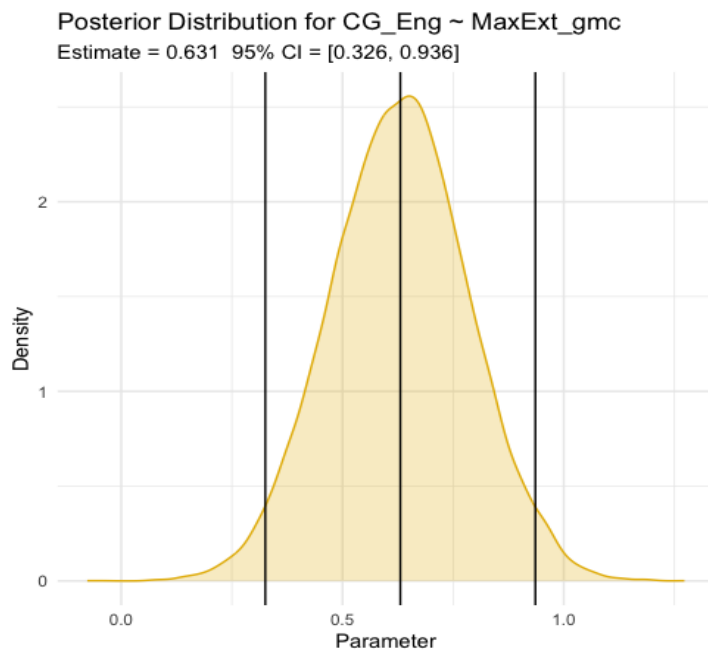


Figure 9

Posterior Distribution of the Effect of Maximum Practice Extensiveness on Caregiver Session-

Level Engagement (Study 2 Aim 2)



Appendix A

BLIMP Syntax

Study 1 Aim 2

```
DATA: Study1Data.csv;
VARIABLES: ;
ORDINAL: ;
NOMINAL: PlanStatus Site StudyCondition;
MISSING: ;
COUNT: TotMissteps;
CLUSTERID: SuperviseeID CaseID;
MODEL: TotMissteps ~ PlanStatus Lag_gmc Site StudyCondition
PlanStatus*StudyCondition;
SEED: 17;
BURN: 30000 ;
ITERATIONS: 30000 ;
```

Study 2 Aim 1

```
DATA: Study2Data_clean.csv;
VARIABLES: ;
ORDINAL: ;
NOMINAL: Site StudyCondition PlanMade;
MISSING: ;
CLUSTERID: SuperviseeID CaseID;
MODEL: AverageExtensiveness ~ TotMissteps TotPrac_gmc Site StudyCondition
PlanMade TotMissteps*PlanMade ;
SEED: 17;
BURN: 30000;
ITERATIONS: 30000;
```

Study 2 Exploratory Analyses

```
DATA: Study1Data.csv;
VARIABLES: ;
ORDINAL: ;
NOMINAL: PlanStatus Site StudyCondition;
MISSING: ;
COUNT: TotMissteps;
CLUSTERID: SuperviseeID CaseID;
MODEL: TotMissteps ~ Lag_gmc PlanStatus MaxExt_gmc Site StudyCondition
PlanStatus*StudyCondition;
SEED: 17;
BURN: 40000 ;
ITERATIONS: 50000 ;
```

Study 2 Aim 2

```
DATA: Study3_YOUTH.csv;
VARIABLES: ;
ORDINAL: ;
NOMINAL: Site StudyCondition;
MISSING: MISSING;
CLUSTERID: SuperviseeID CaseID;
MODEL: Y_Eng ~ TotMissteps MaxExt_gmc Y_MTT_gmc Site StudyCondition;
SEED: 17;
BURN: 20000;
ITERATIONS: 20000;
```

```
DATA: Study3_CAREGIVER.csv;
VARIABLES: ;
ORDINAL: ;
NOMINAL: Site StudyCondition;
MISSING: MISSING;
CLUSTERID: SuperviseeID CaseID;
MODEL: CG_Eng ~ TotMissteps MaxExt_gmc CG_MTT_gmc Site StudyCondition;
SEED: 17;
BURN: 80000;
ITERATIONS: 20000;
```

Appendix B

Equations and Coefficient Interpretations

Study 1 Aim 2

Level 1 Equation (Session):

$$\begin{aligned}\log(\lambda_{ijk}) = & b_{0jk} + b_{1jk}(LAG_{ijk}) + b_{2jk}(NOPLAN_{ijk}) + b_{3jk}(NONENG_{ijk}) \\ & + b_{4jk}(UNMATCH_{ijk}) \\ MISSTEP_{ijk} \sim & NegativeBinomial(\lambda_{ijk}, \phi)\end{aligned}$$

Level-2 Equation (Client):

$$\begin{aligned}b_{0jk} &= \beta_{00k} + \zeta_{0jk} \\ b_{1jk} &= \beta_{10k} \\ b_{2jk} &= \beta_{20k} \\ b_{3jk} &= \beta_{30k} \\ b_{4jk} &= \beta_{40k} \\ \zeta_{0jk} &\sim N(0, \sigma_{client}^2)\end{aligned}$$

Level-3 Equation (Therapist):

$$\begin{aligned}\beta_{00k} &= \gamma_{000} + \gamma_{001}(SITE_k) + \gamma_{002}(CONDITION_k) + v_{00k} \\ \beta_{10k} &= \gamma_{100} \\ \beta_{20k} &= \gamma_{200} + \gamma_{201}(CONDITION_k) \\ \beta_{30k} &= \gamma_{300} + \gamma_{301}(CONDITION_k) \\ \beta_{40k} &= \gamma_{400} + \gamma_{401}(CONDITION_k) \\ v_{00k} &\sim N(0, \sigma_{therapist}^2)\end{aligned}$$

Composite Equation (Full Random Intercept Model):

$$\begin{aligned}\log(\lambda_{ijk}) = & \gamma_{000} + \gamma_{001}(SITE_k) + \gamma_{002}(CONDITION_k) \\ & + \gamma_{100}(LAG_{ijk}) \\ & + [\gamma_{200} + \gamma_{201}(CONDITION_k)](NOPLAN_{ijk}) \\ & + [\gamma_{300} + \gamma_{301}(CONDITION_k)](NONENG_{ijk}) \\ & + [\gamma_{400} + \gamma_{401}(CONDITION_k)](UNMATCH_{ijk}) \\ & + v_{00k} + \zeta_{0jk} \\ MISSTEP_{ijk} \sim & NegativeBinomial(\lambda_{ijk}, \phi)\end{aligned}$$

Study 1 Aim 2 Model Coefficients

Coefficient	Interpretation
γ_{000}	Estimated intercept (i.e., log expected misstep count) when supervision-to-treatment lag is at its mean (i.e., 0), a matched engagement plan was present, and the therapist was in the CKS condition (i.e., 0) and in CA (i.e., 0)
γ_{001}	Effect of site on the intercept—difference in log expected misstep count between a therapist in CA and SC, holding all other covariates constant.
γ_{002}	Effect of study condition on the intercept—difference in log expected misstep occurrences between a therapist in the CKS and PG conditions, holding all other covariates constant.
γ_{100}	Effect of supervision-to-treatment lag on log expected misstep counts, holding all other covariates constant
γ_{200}	Expected difference in log misstep count associated with having no treatment plan (vs. a matched engagement plan) for therapists in the CKS condition, holding all other predictors constant.
γ_{201}	Difference in effect of no treatment plan on log expected misstep count between PG and CKS conditions, holding all other covariates constant
γ_{300}	Expected difference in log misstep count associated with having a non-engagement treatment plan (vs. a matched engagement plan) for therapists in the CKS condition, holding all other covariates constant.
γ_{301}	Difference in effect of a nonengagement treatment plan on log expected misstep count between PG and CKS conditions, holding all other covariates constant
γ_{400}	Expected difference in log misstep count associated with having an unmatched engagement treatment plan (vs. a matched engagement plan) for therapists in the CKS condition, holding all other covariates constant.
γ_{401}	Difference in effect of an unmatched treatment plan on log expected misstep count between PG and CKS conditions, holding all other predictors constant
ζ_{0jk}	Client-level random intercept—deviation of client j 's log expected misstep count from the therapist-level mean, after accounting for fixed effects
u_{00k}	Therapist-level random intercept—deviation of therapist k 's log expected misstep count from the grand mean, after accounting for fixed effects
ϕ	Dispersion parameter for the negative binomial distribution

Study 2 Aim 1

Level-1 Equation (Session)

$$\begin{aligned} AvgExt_{ijk} = & b_{0jk} + b_{1jk}(Misstep_{ijk}) + b_{2jk}(TotPractices_{ijk}) + b_{3jk}(Plan_{ijk}) \\ & + b_{4jk}(Misstep_{ijk} \times Plan_{ijk}) + \varepsilon_{ijk} \\ & \varepsilon_{ijk} \sim iid N(0, \sigma_{\varepsilon}^2) \end{aligned}$$

Level-2 Equation (Client)

$$\begin{aligned} b_{0jk} &= \beta_{00k} + \zeta_{0jk} \\ b_{1jk} &= \beta_{10k} \\ b_{2jk} &= \beta_{20k} \\ b_{3jk} &= \beta_{30k} \\ b_{4jk} &= \beta_{40k} \\ \zeta_{0jk} &\sim N(0, \sigma_{client}^2) \end{aligned}$$

Level-3 Equation (Therapist)

$$\begin{aligned} \beta_{00k} &= \gamma_{000} + \gamma_{001}(Site_k) + \gamma_{002}(StudyCondition_k) + v_{00k} \\ \beta_{10k} &= \gamma_{100} \\ \beta_{20k} &= \gamma_{200} \\ \beta_{30k} &= \gamma_{300} \\ \beta_{40k} &= \gamma_{400} \\ v_{00k} &\sim N(0, \sigma_{therapist}^2) \end{aligned}$$

Composite Equation

$$\begin{aligned} PRAC_{EXT_{ijk}} = & \gamma_{000} + \gamma_{001}(Site_k) + \gamma_{002}(StudyCondition_k) \\ & + \gamma_{100}(Misstep_{ijk}) \\ & + \gamma_{200}(TotPractices_{ijk}) \\ & + \gamma_{300}(Plan_{ijk}) \\ & + \gamma_{400}(Misstep_{ijk} \times Plan_{ijk}) \\ & + v_{00k} + \zeta_{0jk} + \varepsilon_{ijk} \end{aligned}$$

Study 2 Aim 1 Model Coefficients

Coefficient	Interpretation
γ_{000}	Estimated intercept (i.e., expected average practice extensiveness) when no missteps occurred, total practices delivered is at its mean, a treatment plan was present (i.e., 0) and the therapist was in the CKS condition (i.e., 0) and in CA (i.e., 0)
γ_{001}	Effect of site on the intercept—difference in expected average practice extensiveness between therapists in CA and SC, holding all other covariates constant.
γ_{002}	Effect of study condition on the intercept—difference in expected average practice extensiveness between a therapist in the CKS and PG conditions, holding all other covariates constant.
γ_{100}	Effect of misstep count on expected average practice extensiveness, holding all other covariates constant
γ_{200}	Effect of total practices delivered on expected average practice extensiveness, holding all other covariates constant
γ_{300}	Expected difference in expected average practice extensiveness in sessions with no treatment plan (vs. plan present), holding all other covariates constant
γ_{400}	Expected difference in effect of misstep count on average practice extensiveness in sessions with no treatment plan (vs. a treatment plan), holding all other covariates constant
ε_{ijk}	Session-level residual variance
ζ_{0jk}	Client-level residual variance
u_{00k}	Therapist-level residual variance

Study 2 Exploratory Analyses

Level 1 Equation (Session):

$$\begin{aligned}\log(\lambda_{ijk}) = & b_{0jk} + b_{1jk}(LAG_{ijk}) + b_{2jk}(NOPLAN_{ijk}) + b_{3jk}(NONENG_{ijk}) \\ & + b_{4jk}(UNMATCH_{ijk}) + b_{5jk}(MaxExtensiveness_{ijk}) \\ MISSTEP_{ijk} \sim & NegativeBinomial(\lambda_{ijk}, \phi)\end{aligned}$$

Level-2 Equation (Client):

$$\begin{aligned}b_{0jk} &= \beta_{00k} + \zeta_{0jk} \\ b_{1jk} &= \beta_{10k} \\ b_{2jk} &= \beta_{20k} \\ b_{3jk} &= \beta_{30k} \\ b_{4jk} &= \beta_{40k} \\ b_{5jk} &= \beta_{50k} \\ \zeta_{0jk} &\sim N(0, \sigma_{client}^2)\end{aligned}$$

Level-3 Equation (Therapist):

$$\begin{aligned}\beta_{00k} &= \gamma_{000} + \gamma_{001}(SITE_k) + \gamma_{002}(CONDITION_k) + v_{00k} \\ \beta_{10k} &= \gamma_{100} \\ \beta_{20k} &= \gamma_{200} + \gamma_{201}(CONDITION_k) \\ \beta_{30k} &= \gamma_{300} + \gamma_{301}(CONDITION_k) \\ \beta_{40k} &= \gamma_{400} + \gamma_{401}(CONDITION_k) \\ \beta_{50k} &= \gamma_{500} \\ v_{00k} &\sim N(0, \sigma_{therapist}^2)\end{aligned}$$

Composite Equation (Full Random Intercept Model):

$$\begin{aligned}\log(\lambda_{ijk}) = & \gamma_{000} + \gamma_{001}(SITE_k) + \gamma_{002}(CONDITION_k) \\ & + \gamma_{100}(LAG_{ijk}) \\ & + [\gamma_{200} + \gamma_{201}(CONDITION_k)](NOPLAN_{ijk}) \\ & + [\gamma_{300} + \gamma_{301}(CONDITION_k)](NONENG_{ijk}) \\ & + [\gamma_{400} + \gamma_{401}(CONDITION_k)](UNMATCH_{ijk}) \\ & + \gamma_{500}(MaxExtensiveness_{ijk}) \\ & + v_{00k} + \zeta_{0jk} \\ MISSTEP_{ijk} \sim & NegativeBinomial(\lambda_{ijk}, \phi)\end{aligned}$$

Study 2 Exploratory Analyses

Coefficient	Interpretation
γ_{000}	Estimated intercept (i.e., log expected misstep count) when supervision-to-treatment lag and maximum practice extensiveness is at its mean (i.e., 0), a matched engagement plan was present, and the therapist was in the CKS condition (i.e., 0) and in CA (i.e., 0)
γ_{001}	Effect of site on the intercept—difference in log expected misstep count between a therapist in CA and SC, holding all other covariates constant.
γ_{002}	Effect of study condition on the intercept—difference in log expected misstep occurrences between a therapist in the CKS and PG conditions, holding all other covariates constant.
γ_{100}	Effect of supervision-to-treatment lag on log expected misstep counts, holding all other covariates constant
γ_{200}	Expected difference in log misstep count associated with having no treatment plan (vs. a matched engagement plan) for therapists in the CKS condition, holding all other predictors constant.
γ_{201}	Difference in effect of no treatment plan on log expected misstep count between PG and CKS conditions, holding all other covariates constant
γ_{300}	Expected difference in log misstep count associated with having a non-engagement treatment plan (vs. a matched engagement plan) for therapists in the CKS condition, holding all other covariates constant.
γ_{301}	Difference in effect of a nonengagement treatment plan on log expected misstep count between PG and CKS conditions, holding all other covariates constant
γ_{400}	Expected difference in log misstep count associated with having an unmatched engagement treatment plan (vs. a matched engagement plan) for therapists in the CKS condition, holding all other covariates constant.
γ_{401}	Difference in effect of an unmatched treatment plan on log expected misstep count between PG and CKS conditions, holding all other predictors constant
γ_{500}	Effect of maximum practice extensiveness on log expected misstep counts, holding all other covariates constant
ζ_{0jk}	Client-level random intercept—deviation of client j 's log expected misstep count from the therapist-level mean, after accounting for fixed effects
u_{00k}	Therapist-level random intercept—deviation of therapist k 's log expected misstep count from the grand mean, after accounting for fixed effects
ϕ	Dispersion parameter for the negative binomial distribution

Study 2 Aim 2

Level 1 Equation (Session):

$$Y_{ijk} = b_{0jk} + b_{1jk}(\text{Missteps}_{ijk}) + b_{2jk}(\text{MaxExtensiveness}_{ijk}) + \varepsilon_{ijk}$$
$$\varepsilon_{ijk} \sim \text{iid } N(0, \sigma_{\varepsilon}^2)$$

Level 2 Equation (Client):

$$b_{0jk} = \beta_{00k} + \beta_{01k}(\text{BaselineMTT}_{jk}) + \zeta_{0jk}$$
$$b_{1jk} = \beta_{10k}$$
$$b_{2jk} = \beta_{20k}$$
$$\zeta_{0jk} \sim N(0, \sigma_{\text{client}}^2)$$

Level 3 Equation (Therapist):

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{Site}_j) + \gamma_{002}(\text{StudyCondition}_k) + v_{00k}$$
$$\beta_{01k} = \gamma_{010}$$
$$\beta_{10k} = \gamma_{100}$$
$$\beta_{20k} = \gamma_{200}$$
$$v_{00k} \sim N(0, \sigma_{\text{therapist}}^2)$$

Composite Equation (Full Random Intercept Model):

$$Y_{ijk} = \gamma_{000} + \gamma_{001}(\text{Site}_j) + \gamma_{002}(\text{StudyCondition}_k)$$
$$+ \gamma_{010}(\text{BaselineMTT}_{jk})$$
$$+ \gamma_{100}(\text{Missteps}_{ijk})$$
$$+ \gamma_{200}(\text{MaxExtensiveness}_{ijk})$$
$$+ v_{00k} + \zeta_{0jk} + \varepsilon_{ijk}$$

Coefficient	Interpretation
γ_{000}	Estimated intercept (i.e., expected session engagement) when no missteps occurred, maximum practice extensiveness and baseline youth/caregiver engagement challenges were at their mean, and the therapist was in the CKS condition (i.e., 0) and in CA (i.e., 0)
γ_{001}	Effect of site on the intercept—difference in expected session engagement between therapists in CA and SC, holding all other covariates constant.
γ_{002}	Effect of study condition on the intercept—difference in expected session engagement between a therapist in the CKS and PG conditions, holding all other covariates constant.
γ_{010}	Effect of baseline youth/caregiver engagement challenges on the intercept, holding all other covariates constant.
γ_{100}	Effect of misstep count on expected session engagement, holding all other covariates constant
γ_{200}	Effect of maximum practice extensiveness delivered on expected session engagement, holding all other covariates constant
ε_{ijk}	Session-level residual variance
ζ_{0jk}	Client-level residual variance
u_{00k}	Therapist-level residual variance

Appendix C

Originally Proposed Models

Study 2 Aim 1

Level 1 Equation (Session):

$$AvgExt_{ijk} = b_{0jk} + b_{1jk}(Misstep_{ijk}) + b_{2jk}(Plan_{ijk}) + b_{3jk}(Misstep_{ijk} \times Plan_{ijk}) + \varepsilon_{ijk}$$
$$\varepsilon_{ijk} \sim iid N(0, \sigma_{\varepsilon}^2)$$

Level 2 Equation (Client):

$$b_{0jk} = \beta_{00k} + \zeta_{0jk}$$
$$b_{1jk} = \beta_{10k}$$
$$b_{2jk} = \beta_{20k}$$
$$b_{3jk} = \beta_{30k}$$
$$\zeta_{0jk} \sim N(0, \sigma_{client}^2)$$

Level 3 Equation (Therapist):

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(Site_k) + \gamma_{002}(StudyCondition_k) + v_{00k}$$
$$\beta_{10k} = \gamma_{100}$$
$$\beta_{20k} = \gamma_{200}$$
$$\beta_{30k} = \gamma_{300}$$
$$v_{00k} \sim N(0, \sigma_{therapist}^2)$$

Composite Equation:

$$PRAC_{EXT_{ijk}} = \gamma_{000} + \gamma_{001}(Site_k) + \gamma_{002}(StudyCondition_k)$$
$$+ \gamma_{100}(Misstep_{ijk})$$
$$+ \gamma_{200}(Plan_{ijk})$$
$$+ \gamma_{300}(Misstep_{ijk} \times Plan_{ijk})$$
$$+ v_{00k} + \zeta_{0jk} + \varepsilon_{ijk}$$

Study 2 Aim 2

Level 1 Equation (Treatment Event):

$$Y_{ijk} = b_{0jk} + b_{1jk}(\text{Missteps}_{ijk}) + \varepsilon_{ijk}$$
$$\varepsilon_{ijk} \sim \text{iid } N(0, \sigma_{\varepsilon}^2)$$

Level 2 Equation (Client):

$$b_{0jk} = \beta_{00k} + \beta_{01k}(\text{BaselineMTT}_{jk}) + \zeta_{0jk}$$
$$b_{1jk} = \beta_{10k}$$
$$\zeta_{0jk} \sim N(0, \sigma_{client}^2)$$

Level 3 Equation (Therapist):

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{Site}_j) + \gamma_{002}(\text{StudyCondition}_k) + v_{00k}$$
$$\beta_{01k} = \gamma_{010}$$
$$\beta_{10k} = \gamma_{100}$$
$$v_{00k} \sim N(0, \sigma_{therapist}^2)$$

Combined Equation:

$$Y_{ijk} = \gamma_{000} + \gamma_{001}(\text{Site}_j) + \gamma_{002}(\text{StudyCondition}_k)$$
$$+ \gamma_{010}(\text{BaselineMTT}_{jk})$$
$$+ \gamma_{100}(\text{Missteps}_{ijk})$$
$$+ v_{00k} + \zeta_{0jk} + \varepsilon_{ijk}$$

References

- Ackerman, S. J., & Hilsenroth, M. J. (2001). A review of therapist characteristics and techniques negatively impacting the therapeutic alliance. *Psychotherapy: Theory, Research, Practice, Training*, 38(2), 171–185. <https://doi.org/10.1037/0033-3204.38.2.171>
- Alfonsson, S., Lundgren, T., & Andersson, G. (2020). Clinical supervision in cognitive behavior therapy improves therapists' competence: a single-case experimental pilot study. *Cognitive behaviour therapy*, 49(5), 425–438. <https://doi.org/10.1080/16506073.2020.1737571>
- Alfonsson, S., Parling, T., Spännargård, Å., Andersson, G., & Lundgren, T. (2018). The effects of clinical supervision on supervisees and patients in cognitive behavioral therapy: a systematic review. *Cognitive behaviour therapy*, 47(3), 206–228. <https://doi.org/10.1080/16506073.2017.1369559>
- Barnett, M. L., Lau, A. S., Lind, T., Wright, B., Stadnick, N. A., Innes-Gomberg, D., Pesanti, K., & Brookman-Frazee, L. (2020). Caregiver Attendance as a Quality Indicator in the Implementation of Multiple Evidence-Based Practices for Children. *Journal of clinical child and adolescent psychology : the official journal for the Society of Clinical Child and Adolescent Psychology*, American Psychological Association, Division 53, 49(6), 868–882. <https://doi.org/10.1080/15374416.2019.1683851>
- Bearman, S. K., Weisz, J.R., Chorpita, B. F., Hoagwood, K., Ward, A., Ugueto, A. M., & Bernstein, A. (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, 518-529.

- Becker, K. D., & Chorpita, B. F. (2016, August). Enhancing the design of engagement interventions to enhance the public health impact of mental health treatments for youth. In K. Becker (Chair), Extending the reach and impact of science on clinical care for youth and families: Looking for new models for the old challenges. Symposium presented at the 23rd NIMH Conference on Mental Health Services Research: Harnessing Science to Strengthen the Public Health Impact, Bethesda, MD.
- Becker, K. D., & Chorpita, B. F. (2017). *Use of research evidence content application (URECA) coding system* [Measurement instrument, Unpublished instrument].
- Becker, K. D., & Chorpita, B. F. (2023). Future directions in youth and family treatment engagement: Finishing the bridge between science and service. *Journal of Clinical Child and Adolescent Psychology*, 52(2), 284–309. <https://doi.org/10.1080/15374416.2023.2169926>
- 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 and Adolescent Psychology*, 47(1), 1–23. <https://doi.org/10.1080/15374416.2017.1326121>
- Becker, K. D., Dickerson, K., Boustani, M. M., & Chorpita, B. F. (2021). Knowing what to do and when to do it: Mental health professionals and the evidence base for treatment engagement. *Administration and Policy in Mental Health and Mental Health Services Research*, 48(2), 201–218. <https://doi.org/10.1007/s10488-020-01067-6>
- Becker, K. D., Park, A. L., Boustani, M. M., & Chorpita, B. F. (2019). A pilot study to examine the feasibility and acceptability of a coordinated intervention design to address treatment

- engagement challenges in school mental health services. *Journal of school psychology*, 76, 78–88. <https://doi.org/10.1016/j.jsp.2019.07.013>
- Becker, K. D., Wu, E. G., Hukill, A., Brandt, N., & Chorpita, B. F. (2021). How do mental health providers assess treatment engagement of youth and caregivers? *Journal of Child and Family Studies*, 30(10), 2527–2538. <https://doi.org/10.1007/s10826-021-02042-x>
- Binder, J. L., & Strupp, H. H. (1997). Negative process: A recurrently discovered and underestimated facet of therapeutic process and outcome in the individual psychotherapy of adults. *Clinical Psychology: Science and Practice*, 4, 121–139. doi:10.1111/j.1468-2850.1997.tb00105.x
- Bjaastad, J. F., Lillevoll, K., Hoffart, A., Nilsen, T. S., Kjøteroe, J., Prescott, P., Røssberg, J. I., Clark, D. M., Dahl, K., Wang, C. E. A., & Lorentzen, V. (2025). The effect of behavioral rehearsal in the training of clinical psychology students in cognitive therapy for social anxiety disorder: A randomized controlled trial. *Behaviour research and therapy*, 193, 104849. <https://doi.org/10.1016/j.brat.2025.104849>
- Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of management*, 36(4), 1065-1105.
- Bottrill, S., Pistrang, N., Barker, C., & Worrell, M. (2010). The use of therapist self-disclosure: Clinical psychology trainees' experiences. *Psychotherapy Research*, 20(2), 165-180.
- Brock, M. J., Batten, S. V., Walser, R. D., & Robb, H. B. (2015). Recognizing common clinical mistakes in ACT: A quick analysis and call to awareness. *Journal of Contextual Behavioral Science*, 4, 139-143.

- Buckley, P., Karasu, T. B., & Charles, E. (1979). Common mistakes in psychotherapy. *The American journal of psychiatry*, 136(12), 1578–1580.
<https://doi.org/10.1176/ajp.136.12.1578>
- Budge, S. L. (2016). To err is human: An introduction to the special issue on clinical errors. *Psychotherapy*, 53(3), 255–256. <https://doi.org/10.1037/pst0000084>
- Casement, P. (2002). *Learning from our mistakes: Beyond dogma in psychoanalysis and psychotherapy*. New York: Guilford Press
- 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. (2014). Structuring the collaboration of science and service in pursuit of a shared vision. *Journal of Clinical Child and Adolescent Psychology*, 43(2), 323–338. <https://doi.org/10.1080/15374416.2013.828297>
- Chorpita, B. F., Becker, K. D., Park, A. L., Lakind, D., Guan, K., Boustani, M. M., Boyd, M. R., Chu, W., Wu, E. G., & Knudsen, K. S. (2025). Cultivating evidence-based clinical reasoning and action in youth mental health care: The Reaching Families multisite randomized trial. *Journal of Consulting and Clinical Psychology*, 93(2), 65–82.
<https://doi.org/10.1037/ccp0000939>
- 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), 387–393. <https://doi.org/10.1037/a0037746>

- Choy-Brown, M., Baslock, D., Cable, C., Marsalis, S., & Williams, N. J. (2022). In Search of the Common Elements of Clinical Supervision: A Systematic Review. *Administration and policy in mental health*, 49(4), 623–643. <https://doi.org/10.1007/s10488-022-01188-0>
- Cirasola, A., Midgley, N., Muran, J. C., Eubanks, C. F., Hunter, E. B., & Fonagy, P. (2024). Repairing alliance ruptures in psychodynamic psychotherapy with young people: The development of a rational–empirical model to support youth therapists. *Psychotherapy*, 61(1), 68–81. <https://doi.org/10.1037/pst0000514>
- Colby, S. M., Monti, P. M., Barnett, N. P., Rohsenow, D. J., Weissman, K., Spirito, A., Woolard, R. H., & Lewander, W. J. (1998). Brief motivational interviewing in a hospital setting for adolescent smoking: a preliminary study. *Journal of consulting and clinical psychology*, 66(3), 574–578. <https://doi.org/10.1037//0022-006x.66.3.574>
- Colli, A., & Lingardi, V. (2009). The Collaborative Interactions Scale: a new transcript-based method for the assessment of therapeutic alliance ruptures and resolutions in psychotherapy. *Psychotherapy research : journal of the Society for Psychotherapy Research*, 19(6), 718–734. <https://doi.org/10.1080/10503300903121098>
- Colli, A., Gentile, D., Condino, V., & Lingardi, V. (2019). Assessing alliance ruptures and resolutions: Reliability and validity of the Collaborative Interactions Scale-revised version. *Psychotherapy research : journal of the Society for Psychotherapy Research*, 29(3), 279–292. <https://doi.org/10.1080/10503307.2017.1414331>
- Coutinho, J., Ribeiro, E., Hill, C., & Safran, J. (2011). Therapists' and clients' experiences of alliance ruptures: a qualitative study. *Psychotherapy research : journal of the Society for Psychotherapy Research*, 21(5), 525–540. <https://doi.org/10.1080/10503307.2011.587469>

- Creed, T.A., & Kendall, P.C. (2005). Therapist alliance-building behavior within a cognitive-behavioral treatment for anxiety in youth. *Journal of Consulting and Clinical Psychology, 73*, 498.
- Crits-Christoph, P., & Connolly Gibbons, M. B. (2021). Psychotherapy process-outcome research: Advances in understanding causal connections. In M. Barkham, W. Lutz, & L. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 263–295). Wiley.
- Danko C, Garbacz L, & Budd K (2016). Outcomes of Parent–Child Interaction Therapy in an urban community clinic: A comparison of treatment completers and dropouts. *Children and Youth Services Review, 60*, 42–51. 10.1016/j.chldyouth.2015.11.007
- Dawes R. M. (2005). The ethical implications of Paul Meehl's work on comparing clinical versus actuarial prediction methods. *Journal of clinical psychology, 61*(10), 1245–1255.
<https://doi.org/10.1002/jclp.20180>
- Dean, S., Britt, E., Bell, E., Stanley, J., & Collings, S. (2016). Motivational interviewing to enhance adolescent mental health treatment engagement: A randomized clinical trial. *Psychological Medicine, 46*(9), 1961–1969.
- de Haan, A. M., Boon, A. E., de Jong, J. T., Hoeve, M., & Vermeiren, R. R. (2013). A meta-analytic review on treatment dropout in child and adolescent outpatient mental health care. *Clinical Psychology Review, 33*(5), 698–711.
<https://doi.org/10.1016/j.cpr.2013.04.005>
- Dorsey S, Berliner L, Lyon AR, Pullmann MD, & Murray LK (2016). A statewide common elements initiative for children's mental health. *The Journal of Behavioral Health Services & Research, 43*(2), 246–261. 10.1007/s11414-014-9430-y

- Du, H., Ke, Z., Jiang, G., & Huang, S. (2022). The performances of Gelman-Rubin and Geweke's convergence diagnostics of Monte Carlo Markov Chains in Bayesian analysis. *Journal of Behavioral Data Science*, 2(2), Article 2.
<https://doi.org/10.35566/jbds/v2n2/p3>
- Du, H., Keller, B., Alacam, E., & Enders, C. K. (2024). Comparing DIC and WAIC for multilevel models with missing data. *Behavior Research Methods*, 56(4), 2731–2750.
<https://doi.org/10.3758/s13428-023-02231-0>
- Duan, C., Knox, S., & Hill, C. E. (2018). Advice giving in psychotherapy. In E. L. MacGeorge & L. M. Van Swol (Eds.), *The Oxford handbook of advice* (pp. 175–195). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190630188.013.11>
- Duong, M. T., Bruns, E. J., Lee, K., Cox, S., Coifman, J., Mayworm, A., & Lyon, A. R. (2021). Rates of Mental Health Service Utilization by Children and Adolescents in Schools and Other Common Service Settings: A Systematic Review and Meta-Analysis. *Administration and policy in mental health*, 48(3), 420–439.
<https://doi.org/10.1007/s10488-020-01080-9>
- Durham, J. I. (2018). Perceptions of microaggressions: Implications for the mental health and treatment of African American youth. *Journal of Infant, Child & Adolescent Psychotherapy*, 17(1), 52–61. <https://doi.org/10.1080/15289168.2017.1412673>
- Elkin, I., Falconnier, L., Smith, Y., Canada, K. E., Henderson, E., Brown, E. R., & McKay, B. M. (2014). Therapist responsiveness and patient engagement in therapy. *Psychotherapy research : journal of the Society for Psychotherapy Research*, 24(1), 52–66.
<https://doi.org/10.1080/10503307.2013.820855>

- Eubanks, C. F., Lubitz, J., Muran, J. C., & Safran, J. D. (2019). Rupture Resolution Rating System (3RS): Development and validation. *Psychotherapy research : journal of the Society for Psychotherapy Research*, 29(3), 306–319.
<https://doi.org/10.1080/10503307.2018.1552034>
- Eubanks, C.F., Muran, J.C., & Safran, J.D. (2015). Rupture Resolution Rating System (3RS): Manual. 10.13140/2.1.1666.8488
- Firth, N., Saxon, D., Stiles, W. B., & Barkham, M. (2019). Therapist and clinic effects in psychotherapy: A three-level model of outcome variability. *Journal of consulting and clinical psychology*, 87(4), 345–356. <https://doi.org/10.1037/ccp0000388>
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian Data Analysis* (3rd ed.). Chapman and Hall/CRC.
<https://doi.org/10.1201/9780429258411>
- Greaves, G. B. (1988). Common errors in the treatment of multiple personality disorder. *Dissociation: Progress in the Dissociative Disorders*, 1(1), 61–66.
- Guan, K., Kim, R. E., Rodas, N. V., Brown, T. E., Gamarra, J. M., Krull, J. L., & Chorpita, B. F. (2019). Emergent Life Events: An In-Depth Investigation of Characteristics and Provider Responses during Youth Evidence-Based Treatment. *Journal of Clinical Child and Adolescent Psychology*, 48(6), 906–921.
<https://doi.org/10.1080/15374416.2018.1496441>
- Guan, K., Levy, M. C., Kim, R. E., Brown, T. E., Reding, M. E. J., Rith-Najarian, L., Sun, M., Lau, A. S., & Chorpita, B. F. (2017). Managing in-session “surprises:” Provider responses to emergent life events during evidence-based treatment

- implementation. *Administration and Policy in Mental Health and Mental Health Services Research*, 44(2), 164–176. <https://doi.org/10.1007/s10488-015-0692-3>
- Haine-Schlagel, R., & Walsh, N. E. (2015). A review of parent participation engagement in child and family mental health treatment. *Clinical child and family psychology review*, 18(2), 133–150. <https://doi.org/10.1007/s10567-015-0182-x>
- Haine-Schlagel, R., Dickson, K. S., Lind, T., Kim, J. J., May, G. C., Walsh, N. E., Lazarevic, V., Crandal, B. R., & Yeh, M. (2022). Caregiver Participation Engagement in Child Mental Health Prevention Programs: a Systematic Review. *Prevention science : the official journal of the Society for Prevention Research*, 23(2), 321–339. <https://doi.org/10.1007/s11121-021-01303-x>
- Hamilton, A. B., & Finley, E. P. (2019). Qualitative methods in implementation research: An introduction. *Psychiatry research*, 280, 112516. <https://doi.org/10.1016/j.psychres.2019.112516>
- Hanson, J. (2005). Should your lips be zipped? How therapist self-disclosure and non-disclosure affects clients. *Counselling & Psychotherapy Research*, 5(2), 96–104. <https://doi.org/10.1080/17441690500226658>
- Harpaz-Rotem, I., Leslie, D., & Rosenheck, R. A. (2004). Treatment retention among children entering a new episode of mental health care. *Psychiatric Services (Washington, D.C.)*, 55(9), 1022–1028. <https://doi.org/10.1176/appi.ps.55.9.1022>
- Henry, W. P., Schacht, T. E., & Strupp, H. H. (1990). Patient and therapist introject, interpersonal process, and differential psychotherapy outcome. *Journal of Consulting and Clinical Psychology*, 58(6), 768-774. <https://doi.org/10.1037/0022-006X.58.6.768>

- Hill, C. E., Kellems, I. S., Kolchakian, M. R., Wonnell, T. L., Davis, T. L., & Nakayama, E. Y. (2003). The therapist experience of being the target of hostile versus suspected-unasserted client anger: Factors associated with resolution. *Psychotherapy Research*, 13(4), 475–491.
- Hill, C. E., Knox, S., & Duan, C. (2023). Psychotherapist advice, suggestions, recommendations: A research review. *Psychotherapy*, 60(3), 295–305. <https://doi.org/10.1037/pst0000476>
- Hill, C. E., Nutt-Williams, E., Heaton, K. J., Thompson, B. J., & Rhodes, R. H. (1996). Therapist retrospective recall of impasses in long-term psychotherapy: A qualitative analysis. *Journal of Counseling Psychology*, 43(2), 207–217.
- Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy Theory, Research, Practice, Training*, 33(2), 332. <https://doi.org/10.1037/0033-3204.33.2.332>
- Hook, J. N., Farrell, J. E., Davis, D. E., DeBlare, C., Van Tongeren, D. R., & Utsey, S. O. (2016). Cultural humility and racial microaggressions in counseling. *Journal of Counseling Psychology*, 63(3), 269–277. <https://doi.org/10.1037/cou0000114>
- Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge.
- Iovoli, F., Baranowski, M., Sander, L. S., & Rubel, J. A. (2025). Interpersonal Problems and Their Impact on Alliance Ruptures in Psychotherapy. *Clinical psychology & psychotherapy*, 32(2), e70071. <https://doi.org/10.1002/cpp.70071>
- Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical psychology review*, 67, 78–93. <https://doi.org/10.1016/j.cpr.2018.08.004>

- Karver, M. S., De Nadai, A. S., Monahan, M., & Shirk, S. R. (2018). Meta-analysis of the prospective relation between alliance and outcome in child and adolescent psychotherapy. *Psychotherapy (Chicago, Ill.)*, 55(4), 341–355.
<https://doi.org/10.1037/pst0000176>
- Karver, M., Shirk, S., Handelsman, J. B., Fields, S., Crisp, H., Gudmundsen, G., & McMakin, D. (2008). Relationship Processes in Youth Psychotherapy: Measuring Alliance, Alliance-Building Behaviors, and Client Involvement. *Journal of Emotional and Behavioral Disorders*, 16(1), 15-28. <https://doi.org/10.1177/1063426607312536>
- Kazdin, A. E. (1996). Dropping out of child psychotherapy: Issues for research and implications for practice. *Clinical child psychology and psychiatry*, 1(1), 133-156.
- Keller, B. T. (2025). Blimp (Version 4.0.0) [Computer software].
- Keller B (2026). `_rblimp`: Integration of 'Blimp' Software into R_. R package version 1.0.0, <https://CRAN.R-project.org/package=rblimp>.
- Keller, B. T., & Enders, C. K. (2026). Blimp user's guide (Version 4). Retrieved from www.appliedmissingdata.com/blimp
- King, R., & Bickman, L. (2017). Is There a Future for Therapists?. *Administration and policy in mental health*, 44(5), 595–597. <https://doi.org/10.1007/s10488-017-0814-1>
- Kivlighan Jr, D. M., Marmarosh, C. L., & Hilsenroth, M. J. (2014). Client and therapist therapeutic alliance, session evaluation, and client reliable change: A moderated actor–partner interdependence model. *Journal of Counseling Psychology*, 61(1), 15-23.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention

- theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Knox, S., Adrians, N., Everson, E., Hess, S., Hill, C., & Crook-Lyon, R. (2011). Clients' perspectives on therapy termination. *Psychotherapy research : journal of the Society for Psychotherapy Research*, 21(2), 154–167.
<https://doi.org/10.1080/10503307.2010.534509>
- Knox, S., Callender, K. A., Mak, T. W., Skaistis, S., & Knowlton, G. (2022). How Graduate-student or Recent Graduate Psychotherapists Experience and Manage Errors in Psychotherapy. *Counselling Psychology Quarterly*, 35(2), 397–420.
<https://doi.org/10.1080/09515070.2020.1843004>
- Kottler, J. A., & Carlson, J. (2013). *Bad therapy: Master therapists share their worst failures*. Routledge.
- Krause M. (2024). Lessons from ten years of psychotherapy process research. *Psychotherapy research : journal of the Society for Psychotherapy Research*, 34(3), 261–275.
<https://doi.org/10.1080/10503307.2023.2200151>
- Lakind, D., Bradley, W. J., Patel, A., Chorpita, B. F., & Becker, K. D. (2022). A Multidimensional Examination of the Measurement of Treatment Engagement: Implications for Children's Mental Health Services and Research. *Journal of clinical child and adolescent psychology : the official journal for the Society of Clinical Child and Adolescent Psychology*, American Psychological Association, Division 53, 51(4), 453–468. <https://doi.org/10.1080/15374416.2021.1941057>
- Lau, A. S., Lind, T., Cox, J., Motamedi, M., Lui, J. H. L., Chlebowsky, C., Flores, A., Diaz, D., Roesch, S., & Brookman-Frazee, L. (2024). Validating a Pragmatic Measure of

- Evidence-Based Practice (EBP) Delivery: Therapist Reports of EBP Strategy Delivery and Associations with Child Outcome Trajectories. *Administration and policy in mental health*, 51(6), 889–905. <https://doi.org/10.1007/s10488-024-01395-x>
- Lilienfeld S. O. (2007). Psychological Treatments That Cause Harm. *Perspectives on psychological science : a journal of the Association for Psychological Science*, 2(1), 53–70. <https://doi.org/10.1111/j.1745-6916.2007.00029.x>
- Luo, X., & Levendosky, A. A. (2025). A systematic review and integrative framework of psychotherapy microprocess: Linking the science of psychological interventions with the art of moment-to-moment practice. *Clinical psychology review*, 121, 102641. <https://doi.org/10.1016/j.cpr.2025.102641>
- Maaß, U., Kühne, F., Eisert, K., Ghalib, J., & Weck, F. (2026). Transfer of individual supervision goals into the subsequent therapy session: implementation quality, not adherence, is linked to therapeutic skills. *Cognitive behaviour therapy*, 1–16. Advance online publication. <https://doi.org/10.1080/16506073.2026.2635466>
- Meehl, P. E. (1957). When shall we use our heads instead of the formula? *Journal of Counseling Psychology*, 4(4), 268–273. <https://doi.org/10.1037/h0047554>
- McKay, M., McCadam, K., & Gonzales, J. (1996). Addressing the barriers to mental health services for inner city children and their caretakers. *Community Mental Health Journal*, 32(4), 353–361.
- Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). The Guilford Press.
- Mizock, L., & Lundquist, C. (2016). Missteps in psychotherapy with transgender clients: Promoting gender sensitivity in counseling and psychological practice. *Psychology of*

Sexual Orientation and Gender Diversity, 3(2), 148–155.

<https://doi.org/10.1037/sgd0000177>

Moltu, C., Binder, P.-E., & Nielsen, G. H. (2010). Commitment under pressure: Experienced therapists' inner work during difficult therapeutic impasses. *Psychotherapy Research*, 20(3), 309–320.

Monti, P. M., Colby, S. M., Barnett, N. P., Spirito, A., Rohsenow, D. J., Myers, M., Woolard, R., & Lewander, W. (1999). Brief intervention for harm reduction with alcohol-positive older adolescents in a hospital emergency department. *Journal of consulting and clinical psychology*, 67(6), 989–994. <https://doi.org/10.1037//0022-006x.67.6.989>

Muran, J. C., Eubanks, C. F., & Samstag, L. W. (2021). One more time with less jargon: An introduction to "Rupture Repair in Practice". *Journal of clinical psychology*, 77(2), 361–368. <https://doi.org/10.1002/jclp.23105>

Nock, M. K., & Ferriter, C. (2005). Parent management of attendance and adherence in child and adolescent therapy: a conceptual and empirical review. *Clinical child and family psychology review*, 8(2), 149–166. <https://doi.org/10.1007/s10567-005-4753-0>

O'Keeffe, S., Martin, P. & Midgley, N. (2020). When adolescents stop psychological therapy: rupture-repair in the therapeutic alliance and association with therapy ending. *Psychotherapy*, 57(4), pp. 471-490. doi: 10.1037/pst0000279

Owen, J., Tao, K. W., & Drinane, J. M. (2019). Microaggressions: Clinical impact and psychological harm. *Microaggression theory: Influence and implications*, 67-85.

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, 47(3), 344–356.

<https://doi.org/10.1007/s10488-019-00989-0>

Park, H. S., Becker, K. D., & Chorpita, B. F. (2025). Understanding How and When Therapists Misstep: Navigating Engagement Challenges in Youth Mental Health Services. *Research on child and adolescent psychopathology*, 53(5), 669–685.

<https://doi.org/10.1007/s10802-025-01297-y>

Park, H. S., Becker, K. D., Chu, W., Reeder, K., Lechuga, C., Ahuna, J. K., ... Chorpita, B. F. (2025). Identification of Potential Therapist “Missteps” in the Context of Engagement in Public Mental Health Services: Analysis of Digital Recordings and Key Informant Validation. *Psychotherapy Research*, 1–15.

<https://doi.org/10.1080/10503307.2025.2521792>

Paula, C. S., Ziebold, C., Ribeiro, W. S., Pan, P. M., Mari, J. J., Bressan, R., Miguel, E. C., Rohde, L. A., Salum, G. A., & Evans-Lacko, S. (2022). The influence of caregiver attitudes and socioeconomic group on formal and informal mental health service use among youth. *European psychiatry : the journal of the Association of European Psychiatrists*, 65(1), e34. <https://doi.org/10.1192/j.eurpsy.2022.24>

Pellerin, K. A., Costa, N. M., Weems, C. F., & Dalton, R. F. (2010). An examination of treatment completers and non-completers at a child and adolescent community mental health clinic. *Community Mental Health Journal*, 46(3), 273–281.

<https://doi.org/10.1007/s10597-009-9285-5>

Pereira, G. L., Echevarría-Escalante, D., & Froxán-Parga, M. X. (2025). Mapping the dialogue of change: A systematic review of sequential and moment-to-moment dynamics in psychotherapy. *Psychotherapy research : journal of the Society for Psychotherapy*

Research, 1–14. Advance online publication.

<https://doi.org/10.1080/10503307.2025.2565237>

- Piselli, A., Halgin, R. P., & Macewan, G. H. (2011). What went wrong? Therapists' reflections on their role in premature termination. *Psychotherapy research : journal of the Society for Psychotherapy Research*, 21(4), 400–415. <https://doi.org/10.1080/10503307.2011.573819>
- Puhr, R., Heinze, G., Nold, M., Lusa, L., & Geroldinger, A. (2017). Firth's logistic regression with rare events: accurate effect estimates and predictions?. *Statistics in medicine*, 36(14), 2302–2317. <https://doi.org/10.1002/sim.7273>
- Pullmann, M. D., Bruns, E. J., & Sather, A. K. (2013). Evaluating fidelity to the wraparound service model for youth: Application of item response theory to the Wraparound Fidelity Index. *Psychological Assessment*, 25(2), 583–598. <https://doi.org/10.1037/a0031864>
- Reeder, K., Becker, K. D., Chu, W., & Chorpita, B. F. (2024). Characterizing Provider Readiness for Meeting Engagement Challenges in Youth Psychotherapy. *Evidence-Based Practice in Child & Adolescent Mental Health*.
- Renk, K., & Dinger, T. M. (2002). Reasons for therapy termination in a university psychology clinic. *Journal of Clinical Psychology*, 58(9), 1173–1181. <https://doi.org/10.1002/jclp.10075>
- Russell, R., Shirk, S., & Jungbluth, N. (2008). First-session pathways to the working alliance in cognitive-behavioral therapy for adolescent depression. *Psychotherapy Research with Children and Adolescents*, 18, 15–27.
- Rutan, J. S. (2021). Rupture and repair: Using leader errors in psychodynamic group psychotherapy. *International Journal of Group Psychotherapy*, 71(2), 310–331. <https://doi.org/10.1080/00207284.2020.1808471>

- Ryan, R., Berry, K., & Hartley, S. (2023). Review: Therapist factors and their impact on therapeutic alliance and outcomes in child and adolescent mental health - a systematic review. *Child and adolescent mental health*, 28(2), 195–211.
- Safran J.D., Muran J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York, NY: Guilford Press.
- Saloner, B., Carson, N., & Lê Cook, B. (2014). Explaining racial/ethnic differences in adolescent substance abuse treatment completion in the United States: a decomposition analysis. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine*, 54(6), 646–653. <https://doi.org/10.1016/j.jadohealth.2014.01.002>
- Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: theory, research, and practice. *Journal of clinical psychology*, 61(2), 145–153.
<https://doi.org/10.1002/jclp.20107>
- Sarink, F. S. M., & García-Montes, J. M. (2023). Humor interventions in psychotherapy and their effect on levels of depression and anxiety in adult clients, a systematic review. *Frontiers in psychiatry*, 13, 1049476. <https://doi.org/10.3389/fpsyt.2022.1049476>
- Scott, K., & Lewis, C. C. (2015). Using Measurement-Based Care to Enhance Any Treatment. *Cognitive and behavioral practice*, 22(1), 49–59.
<https://doi.org/10.1016/j.cbpra.2014.01.010>
- Sharma, S., & Fowler, J. C. (2016). When countertransference reactions go unexamined due to predetermined clinical tasks: How fear of love can keep us from listening. *Psychotherapy*, 53(3), 302–307.
- Shirk, S. R., & Karver, M. S. (2011). Alliance in child and adolescent psychotherapy. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-based*

- responsiveness* (2nd ed., pp. 70–91). New York, NY: Oxford University Press.
- doi:10.1093/acprof:oso/9780199737208.003.0003
- Shuman, A., & Shapiro, J. (2002). The effects of preparing parents for child psychotherapy on accuracy of expectations and treatment attendance. *Community Mental Health Journal*, 38(1), 3–16. doi:10.1023/A:1013908629870
- Spengler, E. S., Miller, D. J., & Spengler, P. M. (2016). Microaggressions: Clinical errors with sexual minority clients. *Psychotherapy*, 53(3), 360–366. <https://doi.org/10.1037/pst0000073>
- Staudt, M., Lodato, G., & Hickman, C. R. (2012). Therapists talk about the engagement process. *Community Mental Health Journal*, 48(2), 212–218. <https://doi.org/10.1007/s10597-011-9377-x>
- Stige, S. H., Torrisen, M. H., Danielsen, Y. S., Jakobsen, R., Morken, K. T. E., & Dundas, I. (2024). Training students to become responsive therapists: implications from a sequential mixed-methods study on situations that therapists find challenging. *BMC medical education*, 24(1), 261. <https://doi.org/10.1186/s12909-024-05236-1>
- Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: implications for clinical practice. *The American psychologist*, 62(4), 271–286. <https://doi.org/10.1037/0003-066X.62.4.271>
- Taubner, S., Ioannou, Y., Saliba, A., Sales, C. M. D., Volkert, J., Protić, S., Adler, A., Barkauskiene, R., Conejo-Cerón, S., Di Giacomo, D., Mestre, J. M., Moreno-Peral, P., Vieira, F. M., Mota, C. P., Henriques, M. I. R. S., Røssberg, J. I., Perdihi, T. S., Schmidt, S. J., Zettl, M., Ulberg, R., ... Heinonen, E. (2024). Mediators of outcome in adolescent

- psychotherapy and their implications for theories and mechanisms of change: a systematic review. *European child & adolescent psychiatry*, 33(10), 3321–3358.
<https://doi.org/10.1007/s00787-023-02186-9>
- Tchizick, A., Kleinbub, J. R., Bittan, S., Bitton, T., & Zilcha-Mano, S. (2024). Physiological regulation processes differentiate the experience of ruptures between patient and therapist. *Psychotherapy (Chicago, Ill.)*, 61(4), 292–303.
<https://doi.org/10.1037/pst0000543>
- Thompson, R. A., & Sherman, R. T. (1989). Therapist errors in treating eating disorders: Relationship and process. *Psychotherapy: Theory, Research, Practice, Training*, 26(1), 62–68. <https://doi.org/10.1037/h0085406>
- Trimboli, F., Keenan, C. W., & Marshall, R. L. (2016). Common errors in conducting psychodynamic psychotherapy: Illustrative vignettes and alternative strategies. *American Journal of Psychotherapy*, 70(2), 185-202.
- Tsai, K. H., Moskowitz, A. L., Lynch, R. E., Daleiden, E., Mueller, C., Krull, J. L., & Chorpita, B. F. (2019). Do Treatment Plans Matter? Moving From Recommendations to Action. *Journal of clinical child and adolescent psychology : the official journal for the Society of Clinical Child and Adolescent Psychology*, American Psychological Association, Division 53, 48(sup1), S72–S78. <https://doi.org/10.1080/15374416.2016.1204922>
- Valentine, L., & Gabbard, G. O. (2014). Can the use of humor in psychotherapy be taught?. *Academic psychiatry : the journal of the American Association of Directors of Psychiatric Residency Training and the Association for Academic Psychiatry*, 38(1), 75–81. <https://doi.org/10.1007/s40596-013-0018-2>

- van de Schoot, R., & Miočević, M. (2020). Small sample size solutions: A guide for applied researchers and practitioners. Taylor & Francis Group.
- Waehler, C. A., & Grandy, N. M. (2016). Beauty from the beast: Avoiding errors in responding to client questions. *Psychotherapy*, 53(3), 278–283.
- 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>
- Wright, B., Lau, A. S., & Brookman-Frazee, L. (2019). Factors Associated With Caregiver Attendance in Implementation of Multiple Evidence-Based Practices in Youth Mental Health Services. *Psychiatric services (Washington, D.C.)*, 70(9), 808–815.
<https://doi.org/10.1176/appi.ps.201800443>
- Wu, E. G., Becker, K. D., Chu, W., Ahuna, J. K., Lakind, D., Boyd, M. R., ... Chorpita, B. F. (2026). Expanding Integrity Measurement in Mental Health Care: A Coding System of Targets, Practices, and Steps. *Journal of Clinical Child & Adolescent Psychology*, 1–19.
<https://doi.org/10.1080/15374416.2025.2587175>
- Wu, E. G., Becker, K. D., Isenberg, B. M., Park, H. S., Ahuna, J. K., Arkin, S., & Chorpita, B. F. (2026). Gaps in Detection: Comparing Provider- and Client-Reported Indicators of Engagement in Youth Telemental Health. *Journal of Clinical Child & Adolescent Psychology*, 1–14. <https://doi.org/10.1080/15374416.2026.2626905>
- Wu, E. G., Becker, K. D., Kim, R. E., Martinez, J. I., Gamarra, J. M., & Chorpita, B. F. (2022). How Do Treatment Protocols Affect the Use of Engagement Practices in Youth Mental

Health Services?. Administration and policy in mental health, 49(6), 943–961.

<https://doi.org/10.1007/s10488-022-01210-5>

Youngstrom, E. A., Van Meter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, M.-L., &

Youngstrom, J. K. (2017). Evidence-based assessment as an integrative model for

applying psychological science to guide the voyage of treatment. *Clinical Psychology:*

Science and Practice, 24(4), 331–363. <https://doi.org/10.1111/cpsp.12207>

Zlotnick, E., Strauss, A. Y., Ellis, P., Abargil, M., Tishby, O., & Huppert, J. D. (2020).

Reevaluating ruptures and repairs in alliance: Between- and within-session processes in

cognitive-behavioral therapy and short-term psychodynamic psychotherapy. *Journal of*

consulting and clinical psychology, 88(9), 859–869. <https://doi.org/10.1037/ccp0000598>