




# Training Clinical Staff in Genuine CBT: One Large Agency's Preliminary Story

Laurel L. Johnson<sup>1</sup> · Kelli Phythian<sup>1</sup> · Brenda Wong<sup>1</sup> · Claire McGill<sup>1</sup> · Aaliyah Hinds<sup>1</sup> · Ellison Choate<sup>2</sup> · Hannah Zelcer<sup>2</sup> · Eunice Mendez<sup>2</sup> · Yasaman Movahedi<sup>2</sup> · Robert D. Friedberg<sup>2</sup> 

Accepted: 17 July 2023 / Published online: 3 August 2023  
© The Author(s) 2023

## Abstract

While the science documenting CBT's efficacy and effectiveness is strong, workplace applications of the treatment model are often implemented improperly. Therefore, training clinicians in the correct delivery of CBT is essential. This article describes one large agency's initial effort to develop and evaluate a system-wide initiative to supervise clinicians in CBT. Thirty-five clinicians received supervision over 10 sessions and were evaluated on the Cognitive Therapy Rating Scale for Children and Adolescents (CTRS-CA). Client progress was concurrently monitored by the Pediatric Symptom Checklist-17 (PSC-17). The results showed that the initiative was feasible and acceptable. There was an improvement in competency scores ( $t = 4.71, p < 0.001, d = 8.98$ ). Sixty-nine percent of clinicians reached the competency threshold by the end of the training period. Clients also demonstrated significant improvement on the PSC-17 ( $t = 4.31, p < 0.001, d = 4.67$ ). Consequently, this project illustrated the importance of a structured system-wide approach to supervision and training staff to competently deliver CBT.

**Keywords** Cognitive behavioral therapy · Youth · Supervision · Competency · Public mental health agency

## Introduction

Consider this example. You are the director of a large, multi-site children's mental health organization employing a large, diverse staff of practicing child clinicians. The agency is committed to providing evidence-based cognitive behavioral therapy (CBT)

---

✉ Robert D. Friedberg  
rfriedberg@paloaltou.edu

<sup>1</sup> Kinark Child and Family Services, Markham, Ontario, Canada

<sup>2</sup> Center for the Study and Treatment of Anxious Youth, Palo Alto University, Palo Alto, CA, USA

to children, youth, and families with complex needs. While many other agencies and clinics purport to offer such services, you clearly recognize that adherence to genuine CBT is rare and the training outcomes for clients are often questionable. Accordingly, you want to make sure the clinicians within your organization deliver genuine CBT to their clients who desperately need and deserve the highest quality of service. How would you proceed? What would you do to disseminate CBT procedures and train the staff? How would you implement supervision? How would you measure clinicians' competence? This is the story of the way in which one child and youth mental health agency reshaped clinical services and evaluated training outcomes. Our narrative begins with a brief review of the training literature and is followed by an initial evaluation of the agency's competency-based supervision initiative.

CBT spectrum approaches are widely regarded as the gold standard psychosocial interventions for children and adolescents experiencing a myriad of psychological disorders (Davis et al. 2019). Practicing adherent CBT goes beyond following a treatment manual for specific disorders. There are several transdiagnostic core competencies in CBT that promote good treatment outcomes including case conceptualization, assessment, and technical skillfulness in managing inaccurate thoughts, maladaptive behaviors, dysregulated moods, and family dysfunction (Asbaugh et al. 2021; Lusk et al. 2018; Newman 2013; Sburlati et al. 2011). Applying CBT in culturally and developmentally sensitive ways is also required (Friedberg and McClure 2015; Newman 2013). These competencies span a variety of domains such as agenda setting, eliciting feedback, communicating understanding, interpersonal effectiveness, collaboration, pacing, guided discovery, focusing on key cognitions/behaviors, strategies for change, applying techniques, and assigning homework (Lusk et al. 2018; Newman 2013).

CBT training is foundational in most academic settings in fields spanning social work, counseling, psychology, and psychiatry. It is also considered a core competency in specialized training programs such as pre-doctoral psychology internship training and psychiatry residency/fellowship programs. A myriad of professional organizations also offer CBT courses and programs (e.g., The Beck Institute for Cognitive Behavioral Therapy and Research, Academy of Cognitive Therapy, Association of Behavioral and Cognitive Therapies, The Association of Psychological Therapies, etc.). Further, multiple clinician-friendly materials for delivering these procedures to clients are readily available for purchase and as free resources online (Becker et al. 2015; Chorpita and Weisz 2009; Chorpita et al. 2013; Ehrenreich et al. 2018; Weisz and Bearman 2020).

Despite the availability of training opportunities and practice resources, too few young people and their caregivers receive CBT treatment from their mental health service providers (Bearman and Weisz 2015; Williams and Beidas 2019). When left to their own self-evaluation, clinicians tend to overestimate their CBT practice competence (Brosnan 2008; Creed et al. 2016b; Ladany et al. 1996). As many as 71% of clinicians who identified themselves as competent CBT practitioners failed to demonstrate even minimal skillfulness (Creed et al. 2016b). Inaccurately representing oneself as a CBT-oriented provider—in essence “posing” as a cognitive behavioral therapist rather than delivering the appropriate and true approach—harms both clients and the profession (McKay 2014; Sullivan et al. 2014). Consequently,

observation and measurement of trainees' adherence to and competence with CBT practice orthodoxies are imperative (Kazantzis 2003; Kazantzis et al. 2018).

Effective clinical supervision influences both clinicians' competence and clients' therapeutic progress (Falender and Shafranske 2014). Factors that help clinicians achieve competence include participating in structured training courses, regular ongoing supervision, and positive attitudes toward and knowledge of evidence-based treatments (EBTs) (Bearman et al. 2016; Liness et al. 2019). Additionally, deliberate self-reflection and behavioral rehearsal in supervision significantly enhance practice competence (Bearman et al. 2016; Beidas et al. 2014). Moreover, providing supervisees with specific, standardized, constructive, and corrective feedback regarding their application of CBT skills is fundamental to a supervision model that focuses on clinical competency (Bailin et al. 2018; Falender and Shafranske 2014; Milne 2009).

While the research demonstrating the critical factors in clinical supervision is established, Milne and Reiser (2015) reported significant weaknesses within CBT supervision are observed at national, local, and individual levels. More specifically, Schriger et al. (2021) contended that many supervisors tend to neglect evidence-based content in their consultation sessions and deploy little active supervision practices such as supervisory modeling and role play. Observing live sessions or reviewing video and/or audio recordings in supervision is widely recommended to obtain first-hand knowledge of clinical work. Yet, this essential practice is rarely implemented outside of academic settings (Accurso et al. 2011; Bailin et al. 2018). Therefore, concerns about competence in delivering CBT to young clients remain fairly common.

Using CBT adherence rating scales in supervision is a method to evaluate both fidelity and competence among practitioners. Applying these measures in treatment-as-usual community settings is especially important given the potential for greater variability in clinician performance and higher external validity compared to more controlled settings (Goldberg et al. 2020). The Cognitive Therapy Rating Scale (CTRS) (Young and Beck 1980) is regarded as the gold standard for evaluating transdiagnostic CBT competency, adherence, and fidelity with adults (Ashbaugh et al. 2021; Dobson et al. 1985; Creed et al. 2016a, 2021; Goldberg et al. 2020; Muse and McManus 2013; Vallis et al. 1986). As developed by Young and Beck, the CTRS is completed by supervisors after reviewing the trainees' work and covers 11 core elements of CBT practice (agenda setting, homework, key cognitions, etc.). Items are scored on a 7 point scale (0–6) and a score of 40 is the minimum threshold for competence (Goldberg et al. 2020). The instrument enjoys strong psychometric properties (Creed et al. 2016a; Vallis et al. 1986).

Further, supervision and didactic training appear to increase CBT competence levels as measured by the CTRS (Creed et al. 2013). Creed et al. (2013) reported that after 3 months of supervision using this measure, 20% of school-based therapists reached competency benchmarks, and by the 6-month mark, 72% met the criteria for competency. In a subsequent larger and wider study that included 274 clinicians, Creed et al. (2016a) found that 59.5% of clinicians reached competency by the end of their training period. Further, their mixed hierarchical linear model showed that CTRS scores improved over the course of training by on average 18.65 points. Finally, in a recent large-scale study ( $n = 841$ ) of community clinicians, competency ratings

as measured by the CTRS increased over the course of training (Creed et al. 2021). While these results are promising, there are relatively few studies examining the use of rating scales within a community agency-wide CBT competency-based training project conducted in treatment-as-usual contexts. Hence, the quality of both practice and supervision in these naturalistic settings remains somewhat of a black box.

In 2013, the government department that funded children's mental health services in the agency's province released a strategy aimed at improving accessibility and quality of publicly funded mental health care. The agency responded to this directive by implementing evidence-based assessment, intervention, outcome measurement, and supervision practices. Specifically, the agency adopted a competency-based supervision model to support the professional development of clinicians in their evidence-based practices. Falender and Shafanske (2007) defined the competency-based supervision paradigm as an approach that explicitly identifies the knowledge, skills, and values that are assembled to form a clinical competency and develop learning strategies and evaluation procedures to meet criterion-referenced competence standards in keeping with evidence-based practices and the requirements of the local clinical setting (p. 233).

Following Falender and Shafanske's definition, clinical leadership within the agency launched a competency-based model of CBT supervision in 2020 characterized by routine measurement of staff members' skillfulness in CBT with youth.

Despite its many favorable properties, the CTRS suffers from limitations especially when assessing clinician competence in CBT with youth (Affrunti and Creed 2019). Therefore, more child-focused alternative scales were considered for evaluating clinicians' competencies and the Cognitive Therapy Rating Scale for Children and Adolescents was adopted (CTRS-CA; Friedberg and Thordarson 2014; Lusk et al. 2018) for this initiative. To catalyze the accountable practice of CBT, the agency required therapists to submit recordings of treatment sessions to their supervisors and receive ratings using the CTRS-CA.

Accordingly, this study aims to tell the story of this agency's initial effort to increase genuine transdiagnostic CBT competencies among its practitioners working with children and adolescents. Similar to the methods used by Creed et al. (2013, 2016a), the training program's feasibility and acceptability were examined. Additionally, trainee progress was evaluated by analyzing pre-post differences in competency and fidelity ratings. Finally, client progress over the course of the clinician training period was also tracked.

## Method

### Clinical Setting

The clinical setting was among the largest mental health service providers for children and families in a large Canadian province. With an annual budget of nearly \$100 million, the agency provides mental health services to children, adolescents, and families through nine sites spread across roughly 15,000 square kilometers. Approximately 10,000 clients and their families are served annually by the agency's

various locations and services. Clinicians and supervisors who participated in this project cared for diverse youth who presented with a variety of primary psychological disorders as well as comorbid conditions.

## Participants

Thirty-five clinicians participated in this project. In order to protect staff anonymity, no identifying information such as race and gender was recorded. However, similar to other community mental health centers in the province, providers in the agency were predominantly female and White non-Hispanic individuals. Additionally, all therapists were MSW degreed practitioners who had completed an intensive 2-day CBT foundation course and received ongoing weekly supervision by either MSW social workers ( $n = 12$ ) or doctoral level psychologists ( $n = 3$ ). The expected workload for clinicians at the agency is 10–12 clients per week. The average caseload for therapists in the study was 7 clients. Newer hires, less experienced clinicians, and providers who conduct group sessions carry lower numbers of cases. These factors account for the lower caseloads in the study. Clinicians' employment status was not dependent on either their participation in the study or their performance on the competency metrics.

## Instruments

### Cognitive Therapy Rating Scale for Children and Adolescents

Based on the original CTRS, the CTRS-CA (Friedberg and Thordarson 2014) is a 13-item instrument that has been adapted to measure competence in the delivery of CBT to children and adolescents. The CTRS was modified through consultation with experts from the Beck Institute for Cognitive Behavioral Therapy and Research and the Center for Cognitive Therapy at the University of Pennsylvania. The new rating scale retained 10 of the 11 original CTRS items and added 3 unique items (i.e., playfulness, credibility, and informality) that existing literature has revealed to be essential for effective CBT for young patients. Research indicates that playfulness is fundamental to CBT with young people (Chu and Kendall 2009; Peterman et al. 2014; Sburlati et al. 2011; Stallard 2005) and was added as a criterion. Credibility was also included, as confidence in therapists and their abilities have been shown to be highly valued by youth and caregivers (Garcia and Weisz 2002; Shirk and Karver 2006). Therapist informality is another factor valued by patients and their family members (Creed and Kendall 2005) and an item tapping this criterion was also included. Table 1 presents the items that are common and unique to the CTRS and CTRS-CA.

The CTRS-CA is scored on a 7-point scale that ranges from 0 to 6 where higher scores correspond to higher levels of competency. Similar to the CTRS, the CTRS-CA has an accompanying manual (Friedberg 2014) to direct scoring. The competency threshold was set at 52 points, which equates to an average score of 4 across the 13 items. This is roughly equivalent to the 44 point benchmark (i.e., an average score of 4 across 11 items) on the CTRS that was established for certification at the Beck Institute for Cognitive Behavioral Therapy and Research (<https://beckinstitute.org/>).

**Table 1** Comparison of CTRS and CTRS-CA items<sup>1</sup>

CTRS #	CTRS Item	CTRS-CA item	
1	Agenda	7	Agenda
2	Feedback	8	Feedback
3	Understanding	-----	
4	Interpersonal effectiveness	6	Interpersonal effectiveness
5	Collaboration	1	Collaboration
6	Pacing	5	Pacing and pushing
7	Guided discovery	10	Guided discovery
8	Focus on key cognitions	11	Focus on key cognitions
9	Strategy for change	12	Strategy for change/case
10	Use of CBT techniques	13	Use of CBT techniques
11	Homework	9	Homework
	_____	4	Credibility
	_____	2	Informality
	_____	3	Playfulness

<sup>1</sup>CTRS, Cognitive Therapy Rating Scale; CTRS-CA, Cognitive Therapy Rating Scale for Children and Adolescents

CTRS-CA items are classified along three domains: general clinical stance, session structure, and strategies for change. The CTRS-CA has strong internal consistency ( $\alpha = 0.95$ ; Thordarson 2016) which is nearly identical to the original CTRS (Goldberg et al. 2020; Muse and McManus 2013). The measure appears sensitive to change: mean scores have been found to increase over the course of supervision, with the greatest change occurring earlier in the supervision process (Thordarson 2016). Data have also shown that trainees usually enter supervision with better skills in general clinical stance (e.g., collaboration, pacing, and pushing) than in unique CBT competencies (e.g., guided discovery, strategy for change, agenda setting) (Thordarson 2016; Friedberg 2019). Finally, Thordarson (2016) claimed the scale yielded high user satisfaction ratings that indicated ease of scoring as well as consensus that ratings represented accurate estimates of competence.

### Pediatric Symptom Checklist-17 (PSC-17)

The PSC-17 is a widely used brief mental health screener (Gardner et al. 1999, 2007) that assesses changes in psychosocial functioning in children aged 4 to 18 years. A total score is calculated by summing the scores for each of the 17 items that make up the scale and a score that is greater than or equal to 15 suggests a need for referral to a qualified mental health professional. The PSC-17 also contains subscales for internalizing behavior (cutoff  $\geq 5$ ), externalizing behavior (cutoff  $\geq 7$ ), and attention (cutoff  $\geq 7$ ). Children with positive scores on one or more of these subscales are at risk of moderate to severe impairments in emotional or behavioral functioning. Previous studies demonstrate that the PSC-17 is valid, reliable,

and sensitive to change (Chaffin et al. 2017; Murphy et al. 2016). A major advantage of the scale is that it can be completed by caregivers in 3–5 min.

## Procedure

The agency's research review committee, which comprises researchers, subject matter experts, and organizational leaders, reviewed the proposal for scientific merit, feasibility, ethical suitability, and alignment with the agency's strategic priorities. Approval from the agency's research review committee was obtained.

All clients who were receiving treatment at the time of the evaluation and/or their caregivers were invited to participate in the evaluation. They were informed verbally by their therapist of the evaluation and were presented with the informed consent form. Clients were informed that participation in the evaluation was voluntary, they could withdraw at any time, and that the services they received from the agency would not be affected in any way by their decision to participate (or not). Clients and families who agreed to participate signed the consent form. Clients aged 12 and up who had capacity to consent signed on their own behalf. Caregivers of clients under age 12 and of clients aged 12 and up who did not have capacity to consent (as determined by the clinician) signed on their child's behalf. In addition to informed consent, clients and caregivers are informed that data collected by the agency during service may be used for research, evaluation, and/or quality improvement initiatives. This was included in their service agreement, which was signed by clients and/or caregivers upon admission to service.

Staff who were employed as therapists between June and December 2020 were included in the evaluation. Data was collected as part of routine operations and consent was therefore not obtained from participating providers. Justification for not obtaining consent from practitioners, per Tri-Council Policy Statement (TCPS 2 Article 5.5A), was provided to and approved by the community ethics review board. The TCPS 2 is a joint policy of Canada's three federal research agencies—the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Social Sciences and Humanities Research Council of Canada (SSHRC). These agencies require that researchers and their institutions apply the ethical principles and the articles of this policy and be guided by the application sections of the articles. Institutions must ensure that research conducted under their auspices complies with this policy.

Provider names were replaced with numeric codes generated by the investigator. A separate file that connected clinician names with numeric codes exists for internal purposes (i.e., ongoing internal evaluation activities). All data files are password protected and only accessible to research and evaluation staff at the agency. Data was stored on password-protected computers and all files were automatically encrypted. Only aggregated CTRS-CA data was reported (e.g., mean, median, standard deviation, t-score).

Clinicians submitted video recorded sessions of their sessions with clients to their clinical supervisors, who viewed the recordings and scored the CTRS-CA before supervision meetings. Then, in weekly 1-h meetings, supervisors and clinicians reviewed the tapes and CTRS-CA scores, and supervisees were given verbal and



written feedback on their work. CTRS-CA ratings occurred on a weekly basis as part of clinicians' routine supervision meetings. For some clinicians, weekly sessions may have been interrupted by time off (e.g., vacation, sick time); in these circumstances, weekly sessions resumed upon their return to practice. These meetings served as formative evaluations of competency and were used to foster CBT competency and professional development. In addition, supervisors held weekly fidelity huddles where multiple raters evaluated and recorded the same session and compared ratings. Consensus on ratings and divergent scores were both discussed as ways to enhance consistency among raters.

All clinicians completed a mandatory 2-day CBT foundations course prior to the 10-week training period. The course included presentations of CBT theory, the treatment model, session structure, Socratic questioning, guided discovery, basic cognitive behavioral interventions, action plans, and strategies for working with parents. An advanced 2-day course was optional and included instruction in behavioral activation, behavioral experiments, exposure, and intervening with core beliefs.

One week prior to the first session of competency-based supervision, clinicians invited their clients as well as their caregivers to participate in this research. Informed consent was obtained, clients and caregivers were informed that their involvement was voluntary, and that their service would not be affected should they choose not to participate. One week after the clinicians' tenth supervisions session, clients and caregivers who had agreed to the study were once again invited to complete the PSC-17. Completed pre- and post-tests were obtained from a total of 45 clients.

## Results

### Feasibility and Acceptability

Feasibility was assessed by the attrition rate and percentage of supervisees leaving the agency during training (Creed et al. 2013, 2016a). The attrition rate was 1/35 (2.9%) with one trainee dropping out of the program. The very low attrition and turnover rate suggests good feasibility as well as acceptability.

Qualitative feedback gathered from therapists and supervisors after training also supported feasibility and acceptability. The survey results revealed that practitioners were quite satisfied and pleased with their training. Overall, they appreciated and looked forward to more competency-based supervision. Therapists noted an increased sense of self-efficacy with their acquired skills. Receiving timely skill-focused feedback was valued. Technology-related problems (e.g., uploading recordings, how to record, lost or blank audio, etc.) represented a common challenge. Several therapists were concerned about clients' reactions to recording and that the training creating a burdensome workload. However, as previously mentioned, these reservations did not contribute to a high attrition rate. Table 5 lists some specific examples culled from qualitative survey.



**Table 2** Qualitative feedback provided by therapists

Perceived Benefits	Reservations
Pleased with training	Concerned about increased workload
Excited about building CBT skills	Apprehensive about recording
Appreciated the support provided by more structured supervision	Fearful recordings would lead to poor work eval
Looking forward to more constructive feedback	Concern about clients' reactions to recording
Glad supervision is focused on building competency rather than case management	Annoyed by glitches in the technology
Eager to receive both timely and regular Feedback	
Motivated me to acquire new skills	
Felt better about their skill level	

### Inter-rater Reliability

CTRS-CA scores submitted by supervisors during weekly fidelity meetings were used to measure inter-rater reliability for each item (see Table 3). Inter-rater reliability was calculated using Krippendorf's alpha (Krippendorf 2004) because the CTRS-CA yields ordinal rather than continuous data. Further, Krippendorf's alpha ignores missing data as well as having the capacity to handle variability in the number of raters.

**Table 3** Interrater reliability results: Cognitive Therapy Rating Scale for Children and Adolescents (CTRS-CA)

Item	Krippendorf's alpha	95% Confidence interval		Agreement (%)
		Lower bound	Upper bound	
Collaboration	0.636	0.481	0.764	45%
Informality	0.679	0.570	0.812	70%
Playfulness	0.716	0.567	0.753	71%
Credibility	0.541	0.383	0.702	57%
Pacing and pushing	0.513	0.369	0.694	51%
Interpersonal effective	0.615	0.527	0.743	53%
Agenda	0.727	0.670	0.856	61%
Feedback	0.638	0.439	0.752	75%
Homework	0.777	0.725	0.860	60%
Guided discovery	0.572	0.450	0.710	62%
Key cognitions	0.661	0.581	0.781	57%
Strategy for change	0.700	0.616	0.859	69%
Apply CBT	0.736	0.736	0.868	62%

Since this study was a preliminary report on the CTRS-CA, Krippendorff's (2004) recommendation for an acceptable alpha of  $> 0.667$  for tentative investigations was followed. As Table 2 reveals, six items (informality, playfulness, agenda setting, homework assignment, strategy for change, application of CBT techniques) met Krippendorff's criteria for acceptability and four others approached the criterion (collaboration, interpersonal effectiveness, eliciting feedback, focus on key cognitions). Items with the lowest alpha scores included credibility (0.541), pacing and pushing (0.513), and guided discovery (0.572). Percent agreement was also calculated for each item using the formula  $\text{agreement}/(\text{agreement} + \text{disagreement}) \times 100\%$ . Agreement ranged from 45 to 75% with a mean of 61% across the 13 scales.

### Progress on Supervisee Competence and Fidelity Ratings

Descriptive statistics of the CTRS-CA scores at pre- and post-tests are presented in Table 3. On average, scores increased by 7.2 points from pre- to post-test. A paired sample t-test was used to calculate the significance of pre-post differences on supervisees' CTRS-CA scores. Results revealed a significant improvement in scores over the course of ten supervision sessions ( $t = -4.75$ ,  $df = 34$ ,  $p < 0.001$ ; confidence level = 99%tile). Table 4 presents descriptive data for the CTRS-CA scores at pre- and post-test periods. Since this was a small sample, the effect size was also computed and results demonstrated a large effect ( $d = 8.98$ ).

At session 1, the mean score on the CTRS-CA was 46.6 and 20% of clinicians had achieved the threshold for competence (i.e., a score of 52 or higher on at least three assessments). By the end of the 10 sessions, average scores had increased to 53.8, with 69% of clinicians meeting the criterion for competence and 91% scoring above the criterion in at least one session. The average number of sessions needed to meet the competency threshold was 8. Changes in the CTRS-CA are tracked in Fig. 1.

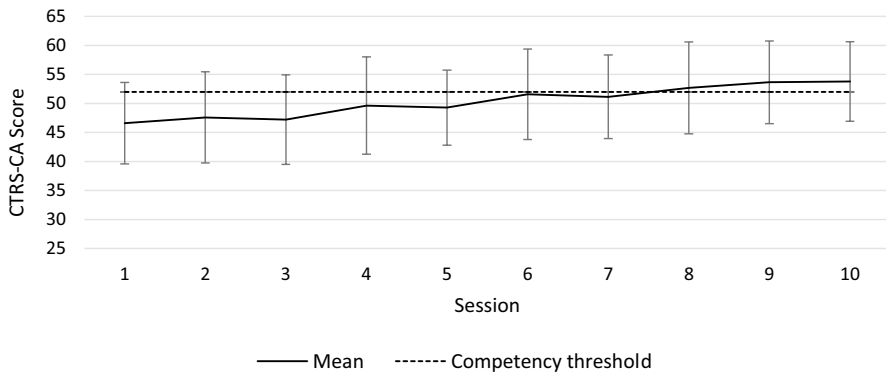
### Client Outcome Data

Including data on client outcomes adds context to the training outcomes. As mentioned above, the clinicians' caseloads represented a wide array of diagnoses and presenting problems. Descriptive analysis of the sample revealed a mean total score of 18 on the PSC-17 at pre-test. The median number of scales on which clients

**Table 4** Cognitive Therapy Rating Scale for Children and Adolescents (CTRS-CA) descriptive results at pre- and post-test

CTRS-CA score	Pre-test	Post-test
Mean	46.6	53.8*
Median	47	53
Minimum	31	41
Maximum	64	69
% greater than 52 (threshold)	20	69

\* $p < 0.001$



**Fig. 1** Change in Mean Cognitive Therapy Rating Scale for Children and Adolescents (CTRS-CA) total scores

exceeded the clinical cutoff was 4, and all participants exceeded the cutoff on at least two scales. Descriptive results are presented in Table 5.

Paired sample t-tests were used to calculate differences in PSC-17 scores from pre- to post-test. Results revealed a statistically significant decline in PSC-17 scores after 10 weeks of supervision (total score:  $t = 4.31$ ,  $df = 44$ ,  $p < 0.001$ ; internalizing subscale:  $t = 3.35$ ,  $df = 44$ ,  $p < 0.01$ ; externalizing subscale:  $t = 3.52$ ,  $df = 44$ ,  $p < 0.01$ ; attention:  $t = 4.77$ ,  $df = 44$ ,  $p < 0.001$ ). Effect sizes ( $d$ ) ranged from 3.22 on the attention scale to 4.67 on the total score, which indicate small to moderate effects. Scores decreased by an average of three points on the total score and about 2 points on each of the subscales.

## Discussion: Lessons Learned

Every good story ends with lessons and this one is no exception. The first takeaway from the study is the implementation of an agency wide competency-based supervision model using objective rating scales was both feasible and acceptable.

**Table 5** Pediatric Symptom Checklist-17 (PSC-17) descriptive results at pre- and post-test for clients treated by supervisees during training

	N	Total score		Internalizing score		Externalizing score		Attention score	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post
Mean	45	18.0	15.0**	11.7	9.9*	12.6	10.7*	11.7	9.4**
SD	45	5.8	6.2	4.7	4.9	3.5	4.4	4.6	4.5
Median	45	17	15	12	9	13	11	10	9
Minimum	45	9	3	1	0	3	2	4	0
Maximum	45	30	30	21	22	22	20	20	19

\*\* $p < 0.001$ ; \* $p < 0.01$

Only one of 35 clinicians dropped out of the project or left the agency during the evaluation period. The low attrition rate compared favorably to data obtained by Creed et al. (2016a) which reported an attrition rate of 17.8%. The feasibility and acceptability are particularly impressive given that the initiative took place during the COVID-19 global pandemic.

Results further revealed that the supervision model contributed to increases in clinician competence. The percentage of therapists reaching competency by the tenth session (69%) compares well to other studies which documented rates of 20% at 3 months (Creed et al. 2013) as well as 60% (Creed et al. 2016a) and 72% (Creed et al. 2013) at final data points. Additionally, improvements in CTRS-CA scores over the course of training/supervision are suggestive of construct validity (Creed et al. 2016a; Cronbach and Meehl 1955; Goldberg et al. 2020). It is important to note that the majority of studies evaluating changes in cognitive behavioral competence measured by rating scales are pre-post designs similar to the one used in the current project (Asbaugh et al. 2021). Therefore, methodological imperfections notwithstanding these initial data suggest the agency-wide competency-based supervision project improved trainees' practices.

Statistically significant improvements in PSC-17 scores from pre- to post-test were observed among clients on each of the four scales. While crude, it is promising that patients being treated by the therapists in this study presented above threshold on all the scales at pre-test and all the scales showed improvement after 10 weeks. Additionally, the young clients cared for by clinicians participating in this study carried multiple diagnoses. Thus, like the recent Creed et al. (2021) study, the community practitioners applied CBT transdiagnostically with very encouraging results.

Several contextual factors likely contributed to the favorable outcomes. First, the supervisors in this study were all experienced, independently licensed clinicians who were trained in CBT, and employed by the agency. This is in contrast with other studies that used post-doctoral fellows or external supervisors to rate competence. Consequently, this study aligns with the recommendation to deploy internal consultants in CBT training efforts to catalyze feasibility, acceptability, and trainee competence (Creed et al. 2016a).

Second, training efforts were ignited by top-down support and driven by the chief executive officer, chief operating officer, and clinical director of the agency. The strong executive and senior leadership support coincided with the agency's mission as well as the government's imperatives for behavioral health care. This commitment from the agency's senior leadership likely enhanced the staff's acceptance of and willingness to participate in the supervision process.

Naturally, clinicians were initially reticent to provide audio or video recordings for review and rating. These feelings of discomfort were validated and normalized, as clinical supervisors were also participating in the supervision process themselves. Additionally, critical infrastructure mechanisms probably reinforced practitioners' involvement. The training architecture included increased clinical supervision in order to ensure solid foundations and expectations for CBT practice, necessary equipment and technology appropriately supplied with ongoing IT support, and importantly, dedicated time provided to clinical supervisors to accommodate the

increased demands by redirecting administrative duties to adjunct services (e.g., human resources, finance, administration). A comprehensive communication plan was established including short, weekly online feedback forms which allowed the clinicians to provide immediate feedback and enabled timely problem solving. Finally, all queries were collated and published as weekly FAQ newsletters.

In terms of future directions, several research areas merit attention. Certainly, randomized clinical trials investigating CBT training initiatives versus various control/contrast conditions are worthy endeavors. However, if the research is done in a non-academic or treatment-as-usual setting, considerable care should be devoted to the control condition. For instance, comparing staff who are randomly assigned to a CBT training course to peers who receive an inactive or no training condition might create some problems in morale and service delivery. Therefore, comparing staff outcomes who receive training to counterparts in a wait-list training condition is a viable option. Recent work by Pachankis and colleagues (2022) offers an interesting design where they randomly assigned staff to a LGBTQ-affirmative CBT training and compared them with providers who were assigned to a 4-month wait list.

Crafting studies which examine specific training structure, content, and delivery modes are also necessary. For instance, comparing on-line to in-person training sessions would likely yield interesting results. Further, assessing whether didactic instruction plus supervision is more effective than either didactic training or clinical supervision alone might illuminate the essential ingredients for implementation of clinical strategies. Testing whether directed or self-guided training produces the best outcomes might help find the most cost-efficient method (Heinrich et al. 2023). Collecting data on providers' pre-existing attitudes regarding CBT would enable moderator studies exploring these variables.

Future projects should consider the clinical and training contexts carefully. Treatment-as-usual settings like the one used in this study have different priorities from academic environments. For instance, limiting costs and building service as well as training capacity are fundamental concerns. Recent training and research program are explicitly addressing issues related to scalability and cost-efficiency (German et al. 2018; Heinrich et al. 2023). Providers, clients, and patients care likely to appreciate research initiatives that respect this clinical context.

## Limitations

The study is a one-group pre-post-test design and, as with other such investigations, is vulnerable to common threats to internal validity (Campbell and Stanley 1963). History, maturation, and testing/repeated measurement effects may have confounded the findings. Additionally, the small  $n$  may also have contributed to type I error. However, the application of Cohen's  $d$  (Cohen 1960) to the data may have mitigated this problem. Finally, supervisors were not blind to the nature of the project and reporter bias may have occurred. Nevertheless, the weekly supervisor meetings and repeated comparisons of CTRS-CA scores among raters may have minimized such bias.

While the inter-rater reliability results were far from perfect, the data appear acceptable. Ten of the 13 items either reached or approached the acceptability

threshold, per Krippendorff's criterion, and the items that are fundamental to CBT practice yielded the highest reliability scores (agenda setting, homework assignment, strategy for change/case conceptualization, and application of techniques). Further, inter-rater reliability has also been a concern in several studies examining the well-established CTRS (Kazantzis et al. 2018; Kuhne et al. 2020; Muse and McManus 2013). In these prior investigations as well as in the present study, multiple factors may have contributed to unwanted heterogeneity in supervisor ratings including the number of recordings that were previously evaluated, raters' training history, variability in the interpretation of anchor points for scoring, supervisors' implicit biases, and supervisors' attitudes toward providing feedback (Kuhne et al. 2020; Muse and McManus 2013). Consequently, future work should consider including these potential factors as co-variates, moderators, and mediators of competency scores.

Previous studies evaluating scores on various cognitive therapy rating scales are criticized for arbitrary competency thresholds (Muse and McManus 2013). Thus, the 52-point score that determined competency could similarly be seen as an incidental benchmark. The CTRS-CA is pending greater psychometric evaluation and more data will better inform decision making.

Finally, more in-depth analysis of the CTRS + CA is recommended. Certainly, larger-scale studies with increased sample sizes which would be more adequately powered represent welcome additions to the literature. Additionally, examining variance in training outcomes as a function of both client and clinician factors is a necessary extension of the current research. Hierarchical linear modeling is a logical choice to determine the impact of predictor variables such as therapist competence or client level of distress (e.g., fixed effects) compared to more random effects.

## Conclusion

Despite the methodological limitations noted above, this agency's story describing a wide-ranging training initiative tells a promising tale. The project demonstrated initial feasibility and acceptability. Additionally, competency scores apparently increased as a function of training and these increased ratings were accompanied by symptom improvement. Future work that examines the durability of these training effects is underway. Ideally, this agency's story provides a compelling case for similar agencies to study the accountability of their CBT training and practice efforts in naturalistic treatment-as-usual settings.

**Funding** Open access funding provided by SCELCL, Statewide California Electronic Library Consortium

## Declarations

**Conflict of Interest** Robert D. Friedberg receives book royalties from Guilford, Springer, John Wiley, Routledge, and Professional Resource Press. He is a consultant to Psychological Assessment Resources and is on the speaking faculty at the Beck Institute for Cognitive Behavior Therapy and Research. All the other authors have nothing to disclose. Portions of this manuscript were presented at the annual meeting of the American Academy of Child and Adolescent Psychiatry.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Accurso, E. C., Taylor, R. M., & Garland, A. F. (2011). Evidence-based practices addressed in community-based children's mental health clinical supervision. *Training and Education in Professional Psychology, 5*, 88–96.
- Affrunti, N. W., & Creed, T. A. (2019). The factor structure of the Cognitive Therapy Rating Scale (CTRS) in a sample of mental health clinicians. *Cognitive Therapy and Research, 43*, 642–655.
- Asbaugh, A. R., Cohen, J. N., & Dobson, K. S. (2021). Training in cognitive behavioural therapy (CBT): National training guidelines from the Canadian Association of Cognitive and Behavioural Therapies. *Canadian Psychology, 62*, 239–251.
- Bailin, A., Bearman, S. K., & Sale, R. (2018). Clinical supervision of mental health professionals serving youth: Form and microskills. *Administration and Policy in Mental Health and Mental Health Research, 45*, 800–812.
- Bearman, S. K., Schneiderman, R. L., & Zoloth, E. (2016). Building an evidence base for effective supervision practices: An analogue experiment of supervision to increase EBT fidelity. *Administration and Policy in Mental Health and Mental Health Services Research, 44*, 293–307.
- Bearman, S. K., & Weisz, J. R. (2015). Review: Comprehensive treatment for youth comorbidity-evidence-guided approaches to a complicated problem. *Child and Adolescent Mental Health, 20*, 131–141.
- Becker, K. D., Lee, B. R., Daleiden, E. L., Lindsey, M., Brandt, N. E., & Chorpita, B. F. (2015). The common elements of engagement in children's mental health services: Which elements for which outcomes? *Journal of Clinical Child and Adolescent Psychology, 44*, 30–43.
- Beidas, R. S., Cross, W., & Dorsey, S. (2014). Show me, don't tell me: Behavioral rehearsal as a training and analogue fidelity tool. *Cognitive and Behavioral Practice, 21*, 1–11.
- Brosnan, L., Reynolds, S., & Moore, R. G. (2008). Self-evaluation of cognitive therapy performance: Do therapists know how competent they are. *Behavioural and Cognitive Therapy, 36*, 581–587.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally.
- Chaffin, M., Campbell, C., Whitworth, D. N., Gillaspay, S. R., Bard, D., Bonner, B. L., & Wolraich, M. L. (2017). Accuracy of a pediatric behavioral health screener to detect untreated behavioral health problems in primary care settings. *Clinical Pediatrics, 56*, 427–434.
- Chorpita, B. F., & Weisz, J. R. (2009). *Modular approach to therapy for children with anxiety, depression, trauma or conduct problems (MATCH-ADTC)*. Practice-Wise.
- Chorpita, B. F., Weisz, J. R., Daleiden, E. L., Schoenwald, S. K., Palinkas, L., Miranda, J. M., Higa-McMillan, C. K., Nakamura, B. J., Austin, A. A., Borntrager, C. F., Ward, A., Wells, K. C., Gibbons, R. D., & the Research Network on Youth Mental Health. (2013). Long-term outcomes for the Child STEPS randomized effectiveness trial: A comparison of modular versus standard treatment design with usual care. *Journal of Consulting and Clinical Psychology, 81*, 999–1009.
- Chu, B. C., & Kendall, P. C. (2009). Therapist responsiveness to child engagement flexibility within fidelity within manual based CBT for anxious children. *Journal of Clinical Psychology, 65*, 736–754.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46.
- Creed, T. A., Crane, M. E., Calloway, A., Olino, T. M., Kendall, P. C., & Stirman, S. W. (2021). Changes in community clinician attitudes and competence in following transdiagnostic CBT training. *Implementation, Research, and Practice, 2*, 1–12.



- Creed, T. A., Frankel, S. A., German, K. L., Jager-Hyman, S., Taylor, K. D., Adler, A. A., Wolk, C. B., Stirman, S. W., Waltman, S. H., Williston, M. A., Sherill, R., Evans, A. C., & Beck, A. T. (2016a). Implementation of transdiagnostic cognitive therapy in community behavioral health: The Beck community initiative. *Journal of Consulting and Clinical Psychology, 84*, 1116–1126.
- Creed, T. A., Jager-Hyman, S., Pontronski, K., Feinberg, B., Rosenberg, Z., Evans, A., Hurford, M. O., & Beck, A. T. (2013). The Beck initiative: Training school-based mental health staff in cognitive therapy. *International Journal of Emotional Education, 5*, 49–66.
- Creed, T. A., & Kendall, P. C. (2005). Therapist alliance building within a cognitive behavioral treatment for anxiety in youth. *Journal of Consulting and Clinical Psychology, 73*, 498–505.
- Creed, T. A., Wolk, C. B., Feinberg, B., Evans, A. C., & Beck, A. T. (2016b). Beyond the label: relationship between community therapists' self-report of a cognitive behavioral therapy orientation and observed Skills. *Administration, Policy, and Research in Mental Health, 43*, 36–43.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281–302.
- Davis, J. P., Palitz, S. A., Knepley, M. J., & Kendall, P. C. (2019). Cognitive-behavioral therapy in youth. In K. S. Dobson & D. J. A. Dozois (Eds.), *Handbook of cognitive behavioral therapies* (4th ed., pp. 349–382). Guilford.
- Dobson, K. S., Shaw, B. F., & Vallis, T. M. (1985). Reliability of a measure of the quality of cognitive therapy. *British Journal of Clinical Psychology, 24*, 295–300.
- Ehrenreich May, J., Kennedy, S. M., Sherman, J. A., Bilek, E. M., Buzzella, B. A., Bennett, S. M., & Barlow, D. H. (2018). *Unified protocols for transdiagnostic treatment of emotional disorders in children and adolescents [therapist guide]*. Oxford University Press.
- Falender, C. A., & Shafranske, E. P. (2007). Competence in competency-based supervision practice: Construct and application. *Professional Psychology: Research and Practice, 38*, 232–240.
- Falender, C. A., & Shafranske, E. P. (2014). Clinical supervision: The state of the art. *Journal of Clinical Psychology: In Session, 70*, 1030–1041.
- Friedberg, R. (2014). *Manual for the cognitive therapy rating scale for children and adolescents (CTRS-CA)*. Center for the Study and Treatment of Anxious Youth at Palo Alto University.
- Friedberg, R. D. (2019). Measuring up: First-session competency ratings on the cognitive therapy rating scale for children and adolescents (CTrs-Ca) for eight practicing clinicians. *Journal of the American Academy of Child & Adolescent Psychiatry, 58*(10), S217.
- Friedberg, R. D., & McClure, J. M. (2015). *Clinical practice of cognitive therapy with children and adolescents* (2nd ed.). Guilford.
- Friedberg, R. D., & Thordarson, M. A. (2014). *The cognitive therapy rating scale for children and adolescents*. Center for the Study and Treatment of Anxious Youth at Palo Alto University.
- Garcia, J. A., & Weisz, J. R. (2002). When youth mental health care stops: Therapeutic relationship problems and other reasons for ending youth outpatient treatment. *Journal of Consulting and Clinical Psychology, 70*, 439–443.
- Gardner, W., Lucas, A., Kolko, D. J., & Campo, J. V. (2007). Comparison of the PSC-17 and alternative mental health screens in an at-risk primary care sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 611–618.
- Gardner, W., Murphy, M., Childs, G., Kelleher, K., Pagano, M., Jellinek, M. S., McInerney, T. K., Wasserman, R. C., Nutting, P., & Chiapetta, L. (1999). The PSC-17: A brief pediatric symptom checklist with psychosocial problem subscales. A report from PROS and ASPN. *Ambulatory Child Health, 5*, 225–236.
- German, R. E., Adler, A., Frankel, S. A., Wiltsey-Stirman, S., Pineda, P., Evans, A. C., Beck, A. T., & Creed, T. A. (2018). Testing a web-based trained peer model to build capacity for evidence-based practices in community mental health systems. *Psychiatric Services, 69*, 286–292.
- Goldberg, S. B., Baldwin, S. A., Merced, K., Caperton, D. D., Imel, Z. E., Atkins, D. C., & Creed, T. A. (2020). The structure of competence: Evaluating the factor structure of the Cognitive Therapy Rating Scale. *Behavior Therapy, 51*, 113–122.
- Heinrich, D., Glombiewski, J. A., & Scholten, S. (2023). Systematic review of training in cognitive-behavior therapy: Summarizing effects, costs, and techniques. *Clinical Psychology Review, 102266*.
- Kazantzis, N. (2003). Therapist competence in cognitive behavioural therapies: review of contemporary empirical evidence. *Behaviour Change, 20*, 1–12.

- Kazantzis, N., Clayton, X., Cronin, T. J., Farchione, D., Limburg, K., & Dobson, K. S. (2018). The cognitive therapy scale and cognitive therapy scale-revised as measures of therapist competence in CBT for depression: Relations with short and long-term outcome. *Cognitive Therapy and Research, 42*, 385–397.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Sage Publications.
- Kuhne, F., Meister, R., MaaB, U., Paunov, T., & Weck, F. (2020). How reliable are therapeutic competence ratings? Results of a systematic review and meta-analysis. *Cognitive Therapy and Research, 44*, 241–247.
- Ladany, N., Hill, C. F., Corbett, M. M., & Nutt, E. A. (1996). Nature, extent, and importance of what psychotherapy trainees do not disclose to their supervisors. *Journal of Counselling Psychology, 43*, 10–24.
- Liness, S., Beale, S., Byrne, S., Hirsch, C. R., & Clark, D. M. (2019). The sustained effects of CBT training on therapist competence and patient outcomes. *Cognitive Therapy and Research, 43*, 631–641.
- Lusk, P., Abney, B. G. H., & Melnyk, B. (2018). A successful model for clinical training in child/adolescent cognitive behavior therapy for graduate psychiatric advanced practice nursing students. *Journal of the American Psychiatric Nurses Association, 24*, 457–468.
- McKay, D. (2014). So you say you are an expert?: False CBT identity harms our hard-earned gains. *The Behavior Therapist, 37*(213), 215–216.
- Milne, D. L. (2009). *Evidence-based clinical supervision: Principles and practice*. Wiley.
- Milne, D. L., & Reiser, R. (2015). Evidence-based supervisory practices in CBT. In D. M. Sudak, R. T. Codd, J. Ludgate, L. Sokol, M. G. Fox, R. Reiser, & D. L. Milne (Eds.), *Teaching and supervising cognitive behavioral therapy* (pp. 207–226). John Wiley.
- Murphy, J. M., Bergmann, P., Chiang, C., Sturmer, R., Howard, B., Abel, M. R., & Jellinek, M. (2016). The PSC-17: Subscale scores, reliability, and factor structure in a new national sample. *Pediatrics, 138*.
- Muse, K., & McManus, F. (2013). A systematic review of methods for assessing competence in cognitive behavioural therapy. *Clinical Psychology Review, 33*, 484–499.
- Newman, C. F. (2013). Training cognitive behavioral supervisors: Didactics, simulated practice, and “meta-supervision”. *Journal of Cognitive Psychotherapy, 27*, 5–16.
- Pachankis, J. E., Soulliard, Z. A., van Dyk, I. S., Layland, E. K., Clark, K. A., Levine, D. S., & Jackson, S. D. (2022). Training in LGBTQ-affirmative cognitive behavior therapy: A randomized clinical trial across LGBTQ community centers. *Journal of Consulting and Clinical Psychology, 90*, 582–599.
- Peterman, J. S., Settiani, C. A., & Kendall, P. C. (2014). Effectively engaging and collaborating with children in cognitive behavioral therapy sessions. In E. S. Sbrurlati, H. J. Lyneham, C. S. Schniering, & R. M. Rapee (Eds.), *Evidence based CBT for anxiety and depression in children and adolescents* (pp. 128–140). Wiley-Blackwell.
- Sbrurlati, E. S., Schniering, C. A., Lyneham, H. J., & Rapee, R. M. (2011). A model of therapist competencies for the empirically supported cognitive behavioral treatment of child and adolescent anxiety and depressive disorders. *Clinical Child and Family Psychology Review, 14*, 89–109.
- Schriger, S. H., Becker-Haimes, E. M., Skriner, L., & Beidas, R., S. (2021). Clinical supervision in community mental health: Characterizing supervision as usual and exploring predictors of supervision content and process. *Community Mental Health Journal, 57*, 552–566.
- Shirk, S. R., & Karver, M. (2006). Process issues in cognitive behavioral therapy for youth. In P. C. Kendall (Ed.), *Child and adolescent therapy* (pp. 465–491). Guilford.
- Stallard, P. (2005). Cognitive behaviour therapy with prepubertal children. In P. Graham (Ed.), *Cognitive behaviour therapy for children and families* (2nd ed., pp. 121–135). Cambridge University Press.
- Sullivan, P., Keller, M., Thordarson, M. A., Trafalis, S., & Friedberg, R. D. (2014). The future is NOW: CBT and health care reform. In N. Columbus (Ed.), *Cognitive Behavioral Intervention* (pp. 1–18). Nova Science.
- Thordarson, M. A. (2016). *The cognitive therapy rating scale for children and adolescents (CTRS-CA): A pilot study*. Doctoral dissertation., Palo Alto University.
- Vallis, T. M., Shaw, B. T., & Dobson, K. S. (1986). The Cognitive therapy scale: Psychometric properties. *Journal of Clinical and Consulting Psychology, 54*, 119–127.

- Weisz, J. R., & Bearman, S. K. (2020). *Principle-guided psychotherapy for children and adolescents: The FIRST program for children and adolescents*. Guilford.
- Williams, N. J., & Beidas, R. S. (2019). Annual research review: The state of implementation science in child psychology and psychiatry: A review and suggestions to advance the field. *Journal of Child Psychology and Psychiatry*, *60*, 430–450.
- Young, J., & Beck, A. T. (1980). *Cognitive therapy scale*. Unpublished manuscript. University of Pennsylvania.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.