

Assessing clinician readiness to deliver exposure therapy: Development and preliminary evaluation of a pragmatic training and implementation tool

Elena L. Pokowitz

pokowitz@umich.edu

University of Michigan

Emily L. Bilek

Michigan Medicine

Research Article

Keywords: exposure therapy, exposure and response prevention, implementation science, measure development

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11380822/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Exposure therapy and exposure and response prevention (EX/RP) are effective but underused treatments for anxiety- and obsessive-compulsive-related disorders. Existing measures assess individual determinants of EX/RP implementation but may be burdensome to administer together. We therefore developed the 28-item Exposure Readiness Scale (ERS), a pragmatic measure assessing readiness to implement EX/RP (through prior training, knowledge, beliefs, confidence, intention, organizational climate), and recent EX/RP use. This quality-improvement project examined the ERS's feasibility, acceptability, and preliminary item and scale performance within the context of a six-hour clinician training program. Across four cohorts, 148 participants initiated 286 ERS administrations. Of these, 95.1% were completed, 3.1% of possible item responses were missing, and median completion time was 5.17 minutes. Participants reported moderate satisfaction and perceived relevance. Qualitative feedback identified relevant caseload, implementation opportunity, skills, and additional barriers as potential areas for refinement. Baseline internal consistency was acceptable for Training ($\alpha = .71$) and strong for Organizational Climate ($\alpha = .87$), but lower for Beliefs ($\alpha = .66$). Most domains demonstrated small-to-moderate positive associations with other domains, suggesting relevance without redundancy. Knowledge showed weak item performance and limited associations with Training and Confidence. The six-domain Readiness score was strongly associated with concurrent self-reported EX/RP Use ($r = .74$, $p < .001$). The ERS is a feasible, promising multidomain implementation tool, but its Knowledge and Beliefs domains, structure, and predictive validity require further refinement and evaluation.

Introduction

Exposure therapy and Exposure and Response Prevention (EX/RP) are considered the gold-standard approach to clinical care in the treatment of avoidance-based disorders such as anxiety, obsessive-compulsive, and other related disorders (Chorpita et al., 2011; Hezel & Simpson, 2019; Higa-McMillan et al., 2016). Despite a strong evidence base, EX/RP is critically underused in mental health care. In fact, EX/RP is one of the least implemented evidence-based practices (EBPs) in community mental health, with less than a third of therapists reporting use with their anxious patients (Becker-Haimes et al., 2020; Chu et al., 2015). This large research-to-practice gap highlights the need for practical methods to determine whether clinicians are prepared to deliver EX/RP and, if not, to identify the supports that they might need to translate training into use with patients.

Readiness to deliver EX/RP is likely shaped by multiple individual- and organization-level factors. At the individual level, clinician knowledge, previous training and supervision, beliefs about exposure, confidence delivering it, and intentions to use it have all previously been identified as important implementation determinants (Becker-Haimes et al., 2017; Frank et al., 2025; Pittig et al., 2019). For example, some providers do not deliver exposure because they believe that exposure is harmful, intolerable, unacceptable to patients, and/or likely to damage the therapeutic alliance (Deacon et al., 2013). Others may have more positive beliefs about exposure but lack confidence in implementing exercises with their patients. Even those providers who are knowledgeable, confident, and motivated to

implement EX/RP might face organization-level factors to exposure delivery, such as inadequate supervision, limited time or resources, lack of support from leadership, or constraints from insurance providers (Becker-Haimes et al., 2017; Becker-Haimes et al., 2020). Readiness to deliver EX/RP therefore reflects multiple concepts beyond just willingness to use the intervention—it involves clinician capability, motivation, and opportunity to do so in their current practicing environment.

This conceptualization of readiness is consistent with the Consolidated Framework for Implementation Research (CFIR 2.0; Damschroder et al., 2022) and the Theoretical Domains Framework (Atkins et al., 2017; Cane et al., 2012), which together highlight determinants that may facilitate or hinder changes in practice. When integrated with extant literature exploring EX/RP specifically, determinants of interest include prior training, knowledge, beliefs, perceived capability, behavioral intentions, and organizational conditions that support the delivery of this EBP (Becker-Haimes et al., 2017; Becker-Haimes et al., 2020; Deacon et al., 2013; Gallo & Barlow, 2012; Lokers, 2020; Racz et al., 2024; Reid et al., 2017, 2018a; Ringle et al., 2015; Scheveneels et al., 2026). Assessing these domains may support trainers and supervisors in identifying why a provider is not yet prepared to use exposure in practice. For example, limited knowledge may indicate a need for enhanced didactics about this approach while low confidence may highlight the opportunity for role plays or consultation.

Several existing instruments assess individual components of readiness to implement EX/RP. The Exposure Training Assessment (ETA; Reid et al., 2018) measures clinicians' previous training and supervision in exposure therapy. The Therapist Beliefs about Exposure Scale (TBES; Deacon et al., 2013) assesses beliefs about the acceptability, tolerability, and safety of exposure, while the Beliefs about Exposure Scale – Training (BES-T; Kemp et al., 2023) probes specifically about clinician perceptions of EX/RP before and after a training program. Knowledge tests and adaptations of the Behavioral Anticipation and Confidence Questionnaire have been used to assess clinician understanding of and confidence in delivering EX/RP (Dimeff et al., 2016; Harned et al., 2013). Other measures are available that address attitudes towards evidence-based practices and organizational conditions in general, such as the Evidence-Based Practice Attitude Scale (EBPAS; Aarons, 2004), the Implementation Climate Scale (ICS; Ehrhart et al., 2014), Implementation Leadership Scale (ILS; Aarons et al., 2014) and pragmatic Context Assessment Tool (pCAT; Robinson & Damschroder, 2023).

Although these instruments provide valuable information, they are likely insufficient when delivered in isolation and burdensome to deliver together, making them infeasible for routine training and clinical settings. General implementation measures may overlook exposure-specific concerns raised in the literature. Further, existing exposure-specific measures typically assess only one domain at a time. Using multiple lengthy measures increases burden for both participants and administrators, particularly when assessments are repeated multiple times to track change. Trainers and clinical programs lack a brief, integrated pragmatic instrument for assessing individual and organization factors most relevant to the delivery of EX/RP. Creating an efficient, multidomain readiness scale could guide the selection of proper supports to increase use of exposure in practice, given that it is understandable, acceptable, brief, and sensitive to change.

The current study

To address this need, we developed the Exposure Readiness Scale (ERS), a brief self-report measure of six proposed determinants of readiness to deliver EX/RP: training, knowledge, beliefs, confidence, intention, and organizational climate. The ERS further includes an assessment of recent use of exposure in practice, theoretically functioning as both an outcome subscale and a continued predictor of future use. This measure was informed by existing exposure-specific and implementation measures, the CFIR 2.0 and TDF, and factors highlighted in the literature as repeatedly being associated with exposure delivery. The ERS was designed to provide provider trainers, supervisors, and clinical programs with a pragmatic profile of clinician- and contextual-level factors that may influence the translation of training into practice.

The present study reports an initial mixed-methods evaluation of the ERS embedded within an experiential exposure-to-exposure training program ([blinded for review]). We examined the measure's feasibility, acceptability, and preliminary item and scale performance. Given the developmental nature of this work, the current study was not intended to establish a definitive factor structure or comprehensive psychometric validity. Rather, we sought to determine whether the ERS is a practical and promising tool to warrant further refinement and evaluation.

Methods

This study was a preliminary mixed-methods evaluation of the ERS conducted in the context of an EX/RP training program delivered by [blinded for review] ([blinded for review]). This evaluation examined measure feasibility, acceptability, and preliminary item and scale performance. The project was conducted as quality assurance/quality improvement (QA/QI) and the [blinded for review, IRB] deemed it non-human-subjects research ([IRB number]). No consent process was therefore required. Participation in data collection (completion of the ERS and other measures) was entirely optional and did not impact eligibility to take part in the training program. Data was collected without direct identifiers and tracked longitudinally via a personally-created identification string.

Participants received the ERS at three timepoints, including pre-training (collected one to three days before the first day of the training), post-training (collected within the week after the second day of the training was completed), and 1-month follow-up (collected between the fourth and fifth week after the second day of the training was completed). Participants in the first two cohorts did not receive any compensation for the completion of their surveys. In cohorts three and four, one \$50 Amazon gift card was raffled off for each timepoint after the training (post- and 1-month) in an attempt to decrease loss to follow-up without a significant change to incentivization. Surveys were administered using Qualtrics. Long-term data management is conducted in REDCap ([blinded for review]).

Exposure-to-exposure training program

Participants completed a six-hour exposure-to-exposure (E2E) training program informed by the learning sciences and designed and delivered by authors [blinded for review]. The training was split into two days (three hours each day). This training was offered for free to mental health care providers over four cohorts: (1) September 2025, [blinded for review], (2) January 2026, sponsored by a university training clinic, (3) May 2026, sponsored by a large university's health and counseling center, and (4) May 2026, sponsored by a rural health provider. The first three trainings were held in [blinded for review]. The fourth training was held in a rural town in [blinded for review].

Participants

Eligible participants were any interested individuals from mental-health-related fields (e.g., psychology, social work, counseling, nursing, etc.) who engaged in at least part of the E2E programs offered between September 2025 and May 2026. Trainees, fully licensed clinicians, and unlicensed staff members were welcome to join. Participants were recruited through the distribution of physical flyers, virtual flyers via email (including relevant Listservs and directed emails), phone call, and fax. One-hundred and fifty-five individuals attended the training across the four cohorts, all of whom were offered the chance to complete the ERS and other measures at pre-training, post-training, and 1-month follow-up.

Measure development

The ERS was developed to provide a brief assessment of factors theorized to influence clinician delivery of EX/RP with patients. Domains were identified through literature review of barriers and facilitators of exposure delivery specifically as well as cross-referencing with the CFIR 2.0 and TDF (Cane et al., 2012; Damschroder et al., 2022). Ultimately, six domains predicting exposure readiness were selected, including (1) prior training and supervision in EX/RP, (2) knowledge of EX/RP, (3) beliefs about EX/RP, (4) confidence in delivering EX/RP, (5) intention to deliver EX/RP, (6) organizational climate regarding EX/RP delivery. A seventh domain, use of EX/RP with patients, was identified as both an important proximal implementation outcome and a potential predictor of continued use of EX/RP.

Items were both adapted from existing instruments and developed by the study team. Existing measures consulted included the ETA (Reid et al., 2018b), Exposure 2 Knowledge and Application Test (Harned et al., 2013), TBES (Deacon et al., 2013), BAQ (Dimeff et al., 2016), BES-T (Kemp et al., 2023), and pCAT (Robinson & Damschroder, 2023). Items adapted from existing measures were used with author permission with the exception of the pCAT, which was created for public adaptation and use. Newly developed items were drafted by authors [blinded for review] and based in literature review regarding both exposure use and the use of EBPs more generally. The initial ERS contained 28 items across seven subscales and was administered without revision across the evaluation period. Contact the primary author for access to the full ERS.

Exposure Readiness Scale (ERS)

Training

The Training subscale consists of four items adapted from the ETA (Reid et al., 2018b). Items assess previous coursework, workshops, or other educational opportunities as well as supervision related to EX/RP both during and after completing their terminal degree. Possible responses range from 0 (none) to 4 (a lot). Two of the questions have a not applicable option coded as 9 (still in graduate training). Higher scores reflect more training and supervision in EX/RP.

Knowledge

The Knowledge subscale contains three items from the Exposure 2 Knowledge and Application Test (Harned et al., 2013). The chosen items assess understanding and application of core principles of EX/RP using multiple-choice questions. Responses are scored 0 (incorrect), 1 (partially correct), or 2 (fully correct). Higher scores reflect more knowledge about EX/RP.

Beliefs

The Beliefs subscale contains seven items. Three of these were adapted from the TBES (Deacon et al., 2013), and four were developed by the study team to assess broader beliefs about psychotherapy that could influence the delivery of EX/RP. Items are rated from 0 (strongly agree) to 4 (strongly disagree). Higher scores reflect more positive beliefs about EX/RP.

Confidence

The Confidence subscale contains three items assessing perceived ability to deliver EX/RP that were informed by the BAQ's prior adaptation for training in exposure (Harned et al., 2013). Responses range from 0 (strongly disagree) to 4 (strongly agree), with higher scores reflecting greater confidence.

Intention

The Intention subscale contains three items, two of which are scored to assess plans and motivation to deliver EX/RP with patients. Responses range from 0 to 4, with anchors varying by item. Higher scores indicate stronger intentions to implement exposure. An additional open-ended item (requesting a numeric response) prompts participants to consider the exact number of patients with whom they plan to incorporate principles of EX/RP within the next week. This item is not included in the subscale score or total score, but rather provides important context for the evaluator to consider regarding readiness. Both scored items on the Intention subscale have a not applicable option (9).

Organizational climate

The Organizational Climate subscale contains six items adapted from the pCAT and informed by the CFIR 2.0 (Damschroder et al., 2022; Robinson & Damschroder, 2023). Items assess access to resources,

supervision, leadership support, and other content relevant to the implementation context of EX/RP. Responses range from 0 (not at all) to 4 (fully), with higher scores indicating an implementation environment more supportive of EX/RP.

Use of EX/RP

Two items assess recent EX/RP delivery in the past month. Items measure the subjective frequency of use of EX/RP with patients as well as the percentage of a provider's overall caseload with which they are integrating components of EX/RP. Responses range from 0 to 4, with a not applicable option (9) for those who did not practice clinical care in the past month. Higher scores reflect more use of EX/RP.

Scoring

Subscale scores were calculated as the mean of completed and eligible items within each domain. Scores on each subscale therefore range from 0 to 4 with the exception of the Knowledge subscale, which ranged from 0 to 2. For the purposes of this development study, "not applicable" (9) options on the Training, Intent, and Use subscales are coded as 0's for scoring.

A preliminary total score was calculated by summing all seven subscale scores, yielding a possible range of 0 to 26. As use of EX/RP is an important implementation outcome, we further delineated a "Readiness score" and "Use score" from the total score on the ERS. The Readiness score is computed through the summation of the six predictive subscales and ranges from 0 to 22, with higher subscale and total scores reflect greater readiness to implement EX/RP. The Use score ranges from 0 to 4 and is simply the average score on the Use subscale.

ERS debriefing survey

Participants completed a debriefing survey regarding the scale itself at the post-training timepoint. This survey included three quantitative items and three qualitative items assessing the ERS's perceived acceptability, relevance, and perceived initial validity. Participants rated their overall satisfaction with the ERS and the extent to which it captured their changes (or lack of changes) in readiness to use EX/RP. Open-ended questions asked participants whether they would want to add or remove any items to the ERS as well as whether they felt the scale captured any changes they experienced related to their readiness. At 1-month follow-up, participants were provided with open space to provide any final comments about the ERS that they felt they did not get the chance to share at post-training.

Demographics

Demographic variables were collected at pre-training. These included basic information about participants and could be skipped or marked "prefer not to answer". Age, gender, race and ethnicity, professional field, primary orientation, current professional status, years in practice, and subjective voluntariness of attending the training program were collected.

Analytic plan

Analyses were conducted using SPSS Version 31 and Excel Version 16. Due to the developmental and exploratory nature of the current study, results are interpreted cautiously.

Acceptability

Acceptability was evaluated using participants' quantitative ratings and qualitative feedback from the ERS debriefing survey. See supplementary materials. Means and standard deviations were calculated for subjective satisfaction and perceived relevance/validity. Original evaluation criteria set by the study team classified average scores of two or higher as reflecting good acceptability, while three or higher would reflect very good acceptability.

Qualitative responses were analyzed using reflexive thematic analysis (Braun & Clarke, 2006, 2019) in a hybrid deductive-inductive approach. Deductive codes reflected the ERS subscale domains, while inductive codes captured concerns or suggestions outside of that scope. All responses were coded by author [blinded for review] in Excel following the RADAR approach to organization and rapid coding and thematization (Watkins, 2017). Author [blinded for review] was the sole coder and was involved in the development and delivery of the E2E training discussed above.

Feasibility

Feasibility was assessed through response rates, item missingness, retention, and completion time. Response rates were measured in four ways, including (1) completed ERS divided by started ERS, (2) unique participants completing the ERS divided by all training participants, (3) completion of the ERS at post-training divided by pre-training, and (4) completion of the ERS at one-month follow-up divided by pre-training. Original evaluation criteria set by the study team classified response rates of 60–79% as moderate and 80%+ as high feasibility.

Overall and item-level missingness were calculated as proportions of possible responses. Study team criteria classified overall item-missingness rates of 5–10% as moderate feasibility and < 5% as high feasibility.

Average completion times were calculated using a built-in feature from Qualtrics allowing time spent on each page of the scale to be documented. Total completion times were calculated as the sum of all time spent on all pages of the ERS, including submission/next page clicks. Average completion times below 10 minutes were deemed moderate feasibility, and those below 5 minutes were deemed high feasibility. Mean, median, and interquartile range were calculated. Completion times over 28 minutes (1 minute per question) were considered extreme outliers.

Preliminary item and scale performance

Item means, standard deviations, ranges, missingness, and response distributions were examined for potential floor or ceiling effects. Corrected item-total correlations were computed within each subscale and within the preliminary total score. Predetermined benchmarks included correlations below 0.30 denoting potentially weak items and those below 0.20 being considered for refinement or replacement.

Cronbach's alpha was calculated for domains containing at least four items, including the Training subscale, Beliefs subscale, and Organizational Climate subscale. Alpha was also calculated for the total Readiness score, including both with and without the open-ended intent item. Values above 0.70 were deemed tentatively supporting internal consistency.

Associations among subscales of the ERS were examined using Spearman rank-order and Pearson correlations. We examined the magnitude and direction of these associations to determine whether domains were theoretically related yet nonredundant. Bootstrapping was used to produce confidence intervals of Pearson correlation estimates.

To avoid treating repeated administrations of the ERS as independent, the primary internal consistency and item analyses use only the baseline timepoint. Analyses using all administrations were considered supplementary.

Readiness as a predictor of use (cross-sectional)

To examine preliminary concurrent criterion-related validity, we evaluated the association between the Readiness score and Use score at baseline. Analyses were restricted to participants who had valid Readiness scores and reported providing clinical care during the previous month ($n = 101$). Pearson and Spearman correlations as well as linear regression were used to estimate this preliminary association. Bootstrapping was used to produce confidence intervals for both correlations.

Results

Participants

The ERS was started 286 times by a total of 148 unique participants. Participants were mostly white, non-Hispanic women. The majority were early career, independently licensed clinicians from the fields of social work and psychology. See Table 1.

Acceptability

Ninety-three participants rated their satisfaction with the ERS at 2.40 out of 4.00 ($sd = 0.89$). Perceived relevance and face validity of the scale was rated by the same participants at 2.33 out of 4.00 ($sd = 0.87$). These quantitative scores were contextualized within qualitative data, which provided overall support for the ERS's acceptability, face validity, and content relevance. Most participants felt that it captured their readiness at least moderately well, and many did not recommend changing its content.

I think that I feel more prepared to do exposure therapy after the training, and I recognized this difference through the survey.

I think the survey captures most or all aspects

Nonetheless, some felt that it inadequately captured pre-post training changes or omitted important dimensions of readiness to implement exposure. The strongest recommendations were to assess relevant caseload or opportunity to implement exposure, to strengthen the knowledge assessment, to add skill-specific readiness items, and to further assess barriers to implementation. Other recommendations (e.g., adding program evaluation questions, making the survey usable for non-clinicians, etc.) were irrelevant to the goals of the current scale.

I think the thing that is still making it difficult is that I am ready/already willing to implement it but my clients are not necessarily willing and it's difficult to build buy in

Include more answer options related to a larger variety of participant scenarios (like more to say I'm a student who has not had clients yet)

Feasibility

Of 286 ERS administrations that were started, 272 (95.10%) completed the prompts in full. 94.70% (143/151) of individuals who took part in the training program fully completed the ERS at least once. Of those who received a valid total ERS score at pre-training, 68.42% (78/114) completed the post-survey and 49.57% (50/114) completed the 1-month follow-up survey.

3.10% of all ERS questions were skipped. See Table 2 for a breakdown of individual question skip rates on the ERS and Table 3 for a breakdown of individual question skip rates on the debriefing survey.

Average completion time was 9.73 minutes. Median completion time was 5.17 minutes, with the middle 50% of respondents finishing between 3.92 and 7.71 minutes (IQR = 3.79 minutes). When outliers over 28 minutes were removed, average completion time was 6.24, median completion time was 5.09, and the middle 50% of respondents finished between 3.88 and 6.93 minutes (IQR = 3.05 minutes).

Preliminary item and scale performance

See Tables 4–7 for item descriptives, including mean, standard deviation, range, percent of valid responses answering the minimum option (labeled minimum %), percent of valid responses answering the maximum option (labeled maximum %), and distribution of responses. See Tables 8 and 9 for corrected item-total correlations within the full scale, the full scale with Intent 1 removed, and within subscales whenever possible (not calculated for Use subscale due to too few questions).

Cronbach's alpha of the full scale (28 items) was 0.90 for 108 valid responses at baseline and 0.89 for 263 valid responses pooled across all timepoints. When Intent 1 was removed (27 items), $\alpha = 0.90$ for 113 valid responses at baseline and $\alpha = 0.91$ for 269 valid responses pooled across all timepoints. When calculated only for those subscales in the Readiness score, $\alpha = 0.88$ for 108 valid responses at baseline and $\alpha = 0.88$ when pooled across all timepoints ($n = 263$). Within the training subscale (4 items), $\alpha = 0.71$ for 122 valid responses at baseline and $\alpha = 0.77$ for 286 valid responses pooled across all timepoints. Within the beliefs subscale (7 items), $\alpha = 0.66$ for 119 valid responses at baseline and $\alpha = 0.78$ for 279 valid responses pooled across all timepoints. Within the organizational climate subscale (6 items), $\alpha =$

0.87 for 114 valid responses at baseline and $\alpha = 0.90$ for 270 valid responses pooled across all timepoints.

Most subscale domains are small to moderately correlated with the other domains with two notable exceptions: the Training and Knowledge subscales (~ 0.17) and the Knowledge and Confidence subscales (~ 0.09). See Table 10 for full domain correlation matrix.

Readiness as a predictor of use (cross-sectional)

The Readiness score at baseline is positively associated with concurrent, self-reported use under both Pearson Correlation, $r(99) = 0.74$, 95% CI [0.65, 0.82], $p < 0.001$, and Spearman's rho, $\rho = 0.75$, 95% CI [0.63, 0.83], $p < 0.001$. The overall model in a simple linear regression was statistically significant, $F(1,99) = 122.94$, $p < .001$, and Readiness accounted for 55.4% of the variance in concurrent Use ($R^2 = .554$, adjusted $R^2 = .549$). A one-point increase in Readiness was associated with a 0.25-point increase in the Use subscale score, $B = 0.248$, $SE = 0.022$, 95% CI [0.203, 0.292], $\beta = .744$, $t(99) = 11.09$, $p < .001$.

The residual histogram and normal probability plot indicated that residuals were approximately normally distributed. The standardized residual by predicted value plot did not indicate a systematic pattern. Two observations had standardized residuals exceeding an absolute value of 3.00; however, neither appeared highly influential based on Cook's distance (maximum = .143). All observations were retained.

Discussion

The current study describes the development and initial mixed-methods evaluation of the Exposure Readiness Scale (ERS), a brief and pragmatic tool for measuring individual- and organization-level factors theorized to influence the delivery of exposure therapy and exposure and response prevention (EX/RP). The initial ERS demonstrates strong administration feasibility, moderate participant-reported acceptability and relevance, and promising but mixed domain performance. It took most participants approximately five minutes to complete. Most proposed readiness domains were positively related but not redundant. The six-domain Readiness score was strongly associated with concurrent self-reported use of EX/RP in practice. The ERS should continue to be revised and tested as a pragmatic implementation tool.

Feasibility and acceptability

Completion rates, item response rates, and median completion times were all strong, suggesting that the ERS can be administered with relatively low burden. This is especially important given the aim of this scale to be used repeatedly in training, consultation, and other clinical settings that require brief and feasible measurement tools. This is distinct, however, from retention of participants longitudinally. A number of participants were lost to follow-up at both post-training and 1-month after the training program, potentially reflecting competing clinical demands, survey fatigue, or issues with the broader

evaluation procedures. Nevertheless, this loss to follow-up suggests that teams wanting to use the ERS longitudinally may benefit from embedding the assessment into the supervision or consultation process, providing protected or paid completion time, or using other retention strategies.

Participants reported moderate satisfaction with the ERS in addition to moderate perceived relevance and face validity. Additional context was provided by qualitative data, which suggested that many participants felt the measure adequately captured their readiness and did not need substantial change. Others, however, noted dimensions they felt were missing (e.g., relevant caseload, opportunity, skill assessment, etc.). This highlights a clear tension between implementation science and scale development, such that additional items or subscales may improve coverage of exposure readiness while also increasing burden and potentially decreasing usefulness. Revisions to the ERS should prioritize items that provide actionable information without substantially extending completion time.

Preliminary item and domain performance

The Training subscale demonstrates acceptable preliminary internal consistency as well as item-domain and item-total correlations. Baseline scores were concentrated towards the low end of the possible responses, which may simply reflect the sample of providers actively seeking training in EX/RP. Performance of this subscale should be examined in samples with broader training and supervision backgrounds in exposure to assess for proper spread of experiences.

The Organizational Climate subscale demonstrated exceptional internal consistency and strong item-domain and item-total correlations. This suggests that a provider's context is relevant to the assessment of readiness, which may otherwise be conceptualized as an individual process. It is important to note that the current items reflect participant perceptions of organizational climate and should not be interpreted as wholly objective; rather, future studies should compare reports within organizations (e.g., trainee vs. supervisor vs. leadership) the extent to which scores are meaningful in assessing the true environment of a clinic or practice. Notably, it is possible that perception of organizational climate may be just as meaningful to an individual provider's readiness to implement exposure as their "true" or objective environment.

Items in the Confidence subscale similarly performed well, with strong item-domain and item-total correlations as well as substantial variation in responses. Intention was strongly correlated with both the Confidence and Use subscales, consistent with research suggesting that clinicians who perceive themselves as capable of exposure delivery are more likely to both intend to do so and to report doing so (Harned et al., 2013; Scheveneels et al., 2026). As the open-ended intention item did not increase or decrease overall reliability of the scale, we recommend retaining this item for optional contextual information.

The Beliefs subscale performed with mixed results, with most items demonstrating modest item-domain and item-total correlations but with baseline internal consistency falling below our predetermined benchmark. It is possible that the domain itself combines exposure-specific beliefs with broader beliefs

about psychotherapy that apply to exposure use. In other words, specific beliefs that exposure is harmful, intolerable, or unethical may be related to, but distinct from, more general beliefs about patient distress and the therapeutic relationship that can apply to other modalities as well. Future research should examine whether the items in this domain map into one coherent factor or two (or more) distinct dimensions. Broader items may still be worth retaining if they seem to contribute to overall readiness in important ways (e.g., in the prediction of exposure use). Additionally, three questions on the scale showed limited range of response. This may be due to wording through a negative lens, anchor options, or asking participants to compare exposure therapy to other modalities in their evaluation of their beliefs. A revised version of the scale may consider revising or removing one or more of the beliefs questions and exploring whether a change in wording may broaden the range of experiences captured.

The Knowledge subscale shows the clearest need for revision, with nearly 90% of participants answering the first item correctly at baseline. This significantly limits the scale's ability to distinguish between clinicians or to track change in knowledge over time. The two other items on this subscale showed increased response variation at baseline but continued to correlate weakly with each other and the broader domain. It is likely that these findings reflect limited content coverage by the length of the subscale rather than irrelevance of knowledge to readiness to deliver exposure. A revision of the ERS would be significantly strengthened by removing, replacing, or increasing the difficulty of the first question. Other revision suggestions include adding one or two brief, case-based questions. As the only objectively scored component of the Readiness score, it is important that the Knowledge subscale not be eliminated solely because it does not function as a more conventional self-report measure.

Domain relationships

Most subscales on the ERS demonstrated small-to-moderate positive associations, which suggests that they are theoretically related without being redundant. These findings are consistent with theoretical guidance of the TDF suggesting that perceived capability, motivation, and opportunity jointly support implementation of an evidence-based practice.

Two domain relationships were notably weak: Training and Knowledge; Knowledge and Confidence. It is possible that the measurement limitations discussed above regarding the Knowledge subscale explain these outlying correlations. However, it is also possible that these findings reflect meaningful distinctions. Participation in exposure-to-exposure training does not inherently ensure retention or application of exposure principles, and factual knowledge on its own does not necessarily produce confidence in conducting exposures with patients. On the other hand, clinicians may report confidence in delivering exposure despite poor knowledge of the underlying principles. These possibilities further highlight the importance of measuring readiness as a cluster of contributing domains rather than through individual factors such as training alone.

While the production of a total Readiness score may be a useful tool for trainers and supervisors, the ERS may also be helpful as a profile of actionable contributors to readiness to implement exposure. Clinicians that have similar overall Readiness scores may have very different needs across the six

domains being measured. In the future, we hope that clinical supervisors find the ERS a useful tool for determining best next steps in training. For example, a low training and/or knowledge score may indicate the need for additional didactic instruction, whereas a low confidence score may signal the need for additional modeling or practice. Supervisors and trainers should examine subscale scores alongside Readiness scores as they work to prepare their trainees for exposure use with patients. Future research should examine whether a true, higher-order factor of Readiness is supported, or whether the ERS should function rather as a more formative index of distinct contributions. Additionally, future research should examine whether a Readiness cutoff can be identified to distinguish between providers who are more or less prepared for delivery of EX/RP. It is further possible that different domains of the Readiness score may contribute at differential weights to “true” readiness, which should be explored.

Readiness as a concurrent predictor of use

It is extremely promising that the six-subscale Readiness score was strongly associated with concurrent self-report use of exposure, accounting for more than half of the variance in Use. This provides preliminary criterion-related evidence that the constructs associated within the ERS are meaningfully related to actual delivery of EX/RP. Sensitivity analyses conducted using Pearson and Spearman estimates further suggest that this finding does not depend upon one correlation method. The magnitude of this association nevertheless warrants thoughtfulness—Readiness and Use were assessed concurrently on the same self-report measure, increasing the possibility of shared-method variance and preventing conclusions about temporal order. Intention, which contributed to the overall Readiness score, was also strongly correlated with the Use subscale. Longitudinal analyses are needed to determine whether the Readiness score predicts later EX/RP use above and beyond use at baseline. These analyses will be reported in the broader training evaluation ([blinded for review]) rather than within this initial performance evaluation of the ERS. Future research would benefit from independent and/or objective indicators of use for comparison, including from sources such as chart review, session coding, or supervisor ratings.

Limitations and future directions

Several limitations qualify the current findings. First, this QA/QI project was embedded within an EX/RP training program rather than an independent psychometric validation study. Participants were largely White, non-Hispanic women working in social work or psychology in the Midwestern United States, and who were often early-career clinicians. The ERS was further evaluated by the same team that developed and delivered both the scale and the training program. The results of this study may not generalize to broader populations, which should be assessed in a formal validation of the ERS.

Second, multidomain Readiness and Use of EX/RP were assessed through self-report and not cross-referenced with objective data. Additionally, exposure delivery claimed on the Use subscale was not verified for accurate or appropriate use with patients. The ERS should not be used to credential providers or to determine competence at this point in its development.

Finally, the current study was unable to establish factor structure, convergent or discriminant validity, test-retest reliability, or predictive validity. Pooled analyses were considered supplementary and exploratory because they included repeated observations. Larger, independent samples will be required to evaluate the underlying structure of the ERS, to compare the ERS with established measures, to examine measurement across broader groups, and to test whether Readiness predicts sustained exposure use over time.

Conclusions

The ERS is a promising brief, integrated assessment of clinician and organizational readiness to deliver exposure therapy. It demonstrates strong administration feasibility, moderate participant acceptability, promising performance across most domains, and a strong concurrent association with self-reported use of exposure. At the same time, weaknesses in the Knowledge and Beliefs domains may have impacted outcomes of the current study. The ERS is currently a promising pragmatic, multidomain profile for training and implementation planning that may in the future become a fully validated measure of readiness and/or competence.

Declarations

Author Contribution

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by EP. The first draft of the manuscript was written by EP and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgement

We would like to thank Dr. Ashley Gearhardt for consulting on this project. We could not have done this work without our undergraduate and post-baccalaureate research assistants, including Imogen Bylinsky, Jacquelyn Wrubel, Sushmitha Peddireddy, Lyra Qian, and Chelsea Coon. Finally, we thank all of the providers who took part in our trainings.

Data Availability

The data that support the findings of this study are currently being used to produce other forthcoming manuscripts and are not available to the public. Individual requests for data sharing should be directed to the corresponding author and will be considered on a case-by-case basis.

References

1. Aarons, G. A. (2004). Mental Health Provider Attitudes Toward Adoption of Evidence-Based Practice: The Evidence-Based Practice Attitude Scale (EBPAS). *Mental Health Services Research*, 6(2), 61–74. <https://doi.org/10.1023/B:MHSR.0000024351.12294.65>
2. Aarons, G. A., Ehrhart, M. G., & Farahnak, L. R. (2014). The implementation leadership scale (ILS): Development of a brief measure of unit level implementation leadership. *Implementation Science*, 9(1), 45. <https://doi.org/10.1186/1748-5908-9-45>
3. Atkins, L., Francis, J., Islam, R., O'Connor, D., Patey, A., Ivers, N., Foy, R., Duncan, E. M., Colquhoun, H., Grimshaw, J. M., Lawton, R., & Michie, S. (2017). A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implementation Science*, 12(1), 77. <https://doi.org/10.1186/s13012-017-0605-9>
4. Becker-Haimes, E. M., Byeon, Y. V., Frank, H. E., Williams, N. J., Kratz, H. E., & Beidas, R. S. (2020). Identifying the organizational innovation-specific capacity needed for exposure therapy. *Depression and Anxiety*, 37(10), 1007–1016. <https://doi.org/10.1002/da.23035>
5. Becker-Haimes, E. M., Okamura, K. H., Wolk, C. B., Rubin, R., Evans, A. C., & Beidas, R. S. (2017). Predictors of clinician use of exposure therapy in community mental health settings. *Journal of Anxiety Disorders*, 49, 88–94. <https://doi.org/10.1016/j.janxdis.2017.04.002>
6. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
7. Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport Exercise and Health*, 11(4), 589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
8. Cane, J., O'Connor, D., & Michie, S. (2012). Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implementation Science*, 7(1), 37. <https://doi.org/10.1186/1748-5908-7-37>
9. Chorpita, B. F., Daleiden, E. L., Ebesutani, C., Young, J., Becker, K. D., Nakamura, B. J., Phillips, L., Ward, A., Lynch, R., Trent, L., Smith, R. L., Okamura, K., & Starace, N. (2011). Evidence-based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice*, 18(2), 154–172. <https://doi.org/10.1111/j.1468-2850.2011.01247.x>
10. Chu, B. C., Crocco, T., Arnold, S., Brown, C. C., Southam-Gerow, R., M. A., & Weisz, J. R. (2015). Sustained implementation of cognitive-behavioral therapy for youth anxiety and depression: Long-term effects of structured training and consultation on therapist practice in the field. *Professional Psychology: Research and Practice*, 46(1), 70–79. <https://doi.org/10.1037/a0038000>
11. Damschroder, L. J., Reardon, C. M., Widerquist, M. A. O., & Lowery, J. (2022). The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation Science*, 17(1), 75. <https://doi.org/10.1186/s13012-022-01245-0>
12. Deacon, B. J., Farrell, N. R., Kemp, J. J., Dixon, L. J., Sy, J. T., Zhang, A. R., & McGrath, P. B. (2013). Assessing therapist reservations about exposure therapy for anxiety disorders: The Therapist Beliefs

- about Exposure Scale. *Journal of Anxiety Disorders*, 27(8), 772–780.
<https://doi.org/10.1016/j.janxdis.2013.04.006>
13. Dimeff, L. A., Koerner, K., Woodcock, E. A., Beadnell, B., Brown, M. Z., Skutch, J. M., Paves, A. P., Bazinet, A., & Harned, M. S. (2016). *Behavioral Anticipation and Confidence Questionnaire* [Dataset].
<https://doi.org/10.1037/t48921-000>
 14. Ehrhart, M. G., Aarons, G. A., & Farahnak, L. R. (2014). Assessing the organizational context for EBP implementation: The development and validity testing of the Implementation Climate Scale (ICS). *Implementation Science*, 9(1), 157. <https://doi.org/10.1186/s13012-014-0157-1>
 15. Frank, H. E., Woodard, G. S., & Martinez, R. G. (2025). Supporting Clinicians in Implementing Exposure Therapy for Anxiety and Related Disorders. *Current Psychiatry Reports*.
<https://doi.org/10.1007/s11920-025-01612-w>
 16. Gallo, K. P., & Barlow, D. H. (2012). Factors Involved in Clinician Adoption and Nonadoption of Evidence-Based Interventions in Mental Health. *Clinical Psychology: Science and Practice*, 19(1), 93–106. <https://doi.org/10.1111/j.1468-2850.2012.01276.x>
 17. Harned, M. S., Dimeff, L. A., Woodcock, E. A., & Contreras, I. (2013). Predicting adoption of exposure therapy in a randomized controlled dissemination trial. *Journal of Anxiety Disorders*, 27(8), 754–762. <https://doi.org/10.1016/j.janxdis.2013.02.006>
 18. Hezel, D., & Simpson, H. (2019). Exposure and response prevention for obsessive-compulsive disorder: A review and new directions. *Indian Journal of Psychiatry*, 61(7), 85.
https://doi.org/10.4103/psychiatry.IndianJPsychiatry_516_18
 19. Higa-McMillan, C. K., Francis, S. E., Rith-Najarian, L., & Chorpita, B. F. (2016). Evidence Base Update: 50 Years of Research on Treatment for Child and Adolescent Anxiety. *Journal of Clinical Child & Adolescent Psychology*, 45(2), 91–113. <https://doi.org/10.1080/15374416.2015.1046177>
 20. Kemp, J., Benito, K., Herren, J., Brown, Z., Frank, H. E., & Freeman, J. (2023). Exposure to exposure: A protocol for leveraging exposure principles during training to address therapist-level barriers to exposure implementation. *Frontiers in Psychiatry*, 14, 1096259.
<https://doi.org/10.3389/fpsy.2023.1096259>
 21. Lokers, L. M. (2020). Fear Over Facts: When Fellow Clinicians Become a Barrier to Dissemination of Exposure and Response Prevention for Obsessive-Compulsive Disorder—A Clinical Response to Licensing Board Investigation of Exposure and Response Prevention. *Journal of Cognitive Psychotherapy*, 34(3), 179–184. <https://doi.org/10.1891/JCPSY-D-20-00005>
 22. Pittig, A., Kotter, R., & Hoyer, J. (2019). The Struggle of Behavioral Therapists With Exposure: Self-Reported Practicability, Negative Beliefs, and Therapist Distress About Exposure-Based Interventions. *Behavior Therapy*, 50(2), 353–366. <https://doi.org/10.1016/j.beth.2018.07.003>
 23. [blinded for review].
 24. [blinded for review].
 25. Racz, J. I., Bialocerkowski, A., Calteaux, I., & Farrell, L. J. (2024). Determinants of Exposure Therapy Implementation in Clinical Practice for the Treatment of Anxiety, OCD, and PTSD: A Systematic

- Review. *Clinical Child and Family Psychology Review*, 27(2), 317–341.
<https://doi.org/10.1007/s10567-024-00478-3>
26. Reid, A. M., Bolshakova, M. I., Guzick, A. G., Fernandez, A. G., Striley, C. W., Geffken, G. R., & McNamara, J. P. (2017). Common Barriers to the Dissemination of Exposure Therapy for Youth with Anxiety Disorders. *Community Mental Health Journal*, 53(4), 432–437.
<https://doi.org/10.1007/s10597-017-0108-9>
27. Reid, A. M., Guzick, A. G., Fernandez, A. G., Deacon, B., McNamara, J. P. H., Geffken, G. R., McCarty, R., & Striley, C. W. (2018a). Exposure therapy for youth with anxiety: Utilization rates and predictors of implementation in a sample of practicing clinicians from across the United States. *Journal of Anxiety Disorders*, 58, 8–17. <https://doi.org/10.1016/j.janxdis.2018.06.002>
28. Reid, A. M., Guzick, A. G., Fernandez, A. G., Deacon, B., McNamara, J. P. H., Geffken, G. R., McCarty, R., & Striley, C. W. (2018b). *Exposure Training Assessment* [Dataset]. <https://doi.org/10.1037/t76219-000>
29. Ringle, V. A., Read, K. L., Edmunds, J. M., Brodman, D. M., Kendall, P. C., Barg, F., & Beidas, R. S. (2015). Barriers to and Facilitators in the Implementation of Cognitive-Behavioral Therapy for Youth Anxiety in the Community. *Psychiatric Services*, 66(9), 938–945.
<https://doi.org/10.1176/appi.ps.201400134>
30. Robinson, C. H., & Damschroder, L. J. (2023). A pragmatic context assessment tool (pCAT): Using a Think Aloud method to develop an assessment of contextual barriers to change. *Implementation Science Communications*, 4(1), 3. <https://doi.org/10.1186/s43058-022-00380-5>
31. Scheveneels, S., Hermans, D., Janssens, J., Greeven, A., & Van Emmerik, A. (2026). The use of exposure in anxiety-related disorders and associations with practitioner characteristics. *Cognitive Behaviour Therapy*, 55(5), 652–667. <https://doi.org/10.1080/16506073.2025.2518424>
32. Watkins, D. C. (2017). The RADaR Technique for Applied Research. *International Journal of Qualitative Methods*, 16(1), 1609406917712131. <https://doi.org/10.1177/1609406917712131>. Rapid and Rigorous Qualitative Data Analysis:.

Tables

Table 1 to 10 are available in the Supplementary Files section.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table110.docx](#)