



**U.S. Department of Health and Human Services
Assistant Secretary for Planning and Evaluation
Office of Disability, Aging and Long-Term Care Policy**

DEVELOPMENT OF A QUALITY MEASURE FOR ADULTS WITH POST-TRAUMATIC STRESS DISORDER

April 2019

Office of the Assistant Secretary for Planning and Evaluation

The Office of the Assistant Secretary for Planning and Evaluation (ASPE) is the principal advisor to the Secretary of the Department of Health and Human Services (HHS) on policy development issues, and is responsible for major activities in the areas of legislative and budget development, strategic planning, policy research and evaluation, and economic analysis.

ASPE develops or reviews issues from the viewpoint of the Secretary, providing a perspective that is broader in scope than the specific focus of the various operating agencies. ASPE also works closely with the HHS operating agencies. It assists these agencies in developing policies, and planning policy research, evaluation and data collection within broad HHS and administration initiatives. ASPE often serves a coordinating role for crosscutting policy and administrative activities.

ASPE plans and conducts evaluations and research--both in-house and through support of projects by external researchers--of current and proposed programs and topics of particular interest to the Secretary, the Administration and the Congress.

Office of Disability, Aging and Long-Term Care Policy

The Office of Disability, Aging and Long-Term Care Policy (DALTCP), within ASPE, is responsible for the development, coordination, analysis, research and evaluation of HHS policies and programs which support the independence, health and long-term care of persons with disabilities--children, working aging adults, and older persons. DALTCP is also responsible for policy coordination and research to promote the economic and social well-being of the elderly.

In particular, DALTCP addresses policies concerning: nursing home and community-based services, informal caregiving, the integration of acute and long-term care, Medicare post-acute services and home care, managed care for people with disabilities, long-term rehabilitation services, children's disability, and linkages between employment and health policies. These activities are carried out through policy planning, policy and program analysis, regulatory reviews, formulation of legislative proposals, policy research, evaluation and data planning.

This report was prepared under contract #HHSP23320100019WI between HHS's ASPE/DALTCP and Mathematica Policy Research. For additional information about this subject, you can visit the DALTCP home page at <https://aspe.hhs.gov/office-disability-aging-and-long-term-care-policy-daltcp> or contact the ASPE Project Officers, D.E.B. Potter, Joel Dubenitz, and Laurel Fuller, at HHS/ASPE/DALTCP, Room 424E, H.H. Humphrey Building, 200 Independence Avenue, S.W., Washington, D.C. 20201; Joel.Dubenitz@hhs.gov.

The opinions and views expressed in this report are those of the authors. They do not reflect the views of the Department of Health and Human Services, the contractor or any other funding organization. This report was completed and submitted on April 30, 2015.

REPORT

FINAL REPORT

The Development of a Quality Measure for Adults with Post-Traumatic Stress Disorder

April 2019

Melissa Azur
Daniel Friend
Dmitriy Poznyak
Kathleen Feeney
Danielle Chelminsky

Breanna Miller
Lareina La Flair
Junqing Liu

Submitted to:
Department of Health and Human Services
Office of the Assistant Secretary for Planning and Evaluation
200 Independence Ave., SW
Humphrey Building, HHH-4 424E.11
Washington, DC 20201
Project Officer: Kirsten Beronio
Contract Number: HHSP23320100019WI

Submitted by:
Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Melissa Azur
Reference Number: 06986

TABLE OF CONTENTS

ACKNOWLEDGMENTS	v
ABSTRACT	vi
ACRONYMS	vii
EXECUTIVE SUMMARY	ix
I. PROJECT RATIONALE	1
A. Project Purpose.....	1
B. Report Roadmap.....	2
II. SELECTION OF MEASURE CONCEPTS	3
A. Environmental Scan of PTSD Treatments and Measures	3
B. Scan of Measures.....	4
C. Technical Advisory Group Review	4
D. Selection of Measure Concept.....	5
III. MEASURE SPECIFICATION	8
A. Selection of Data Source	8
B. Identification of Critical PTSD Psychotherapy Treatment Elements using an Established Methodology	8
C. Survey Item Development.....	9
IV. APPROACH TO MEASURE TESTING.....	11
A. Stage 1--Testing the Survey Items	11
B. Stage 2--Pre-Testing the Measure	13
V. TESTING RESULTS.....	27
A. Summary of Survey Administration.....	27
B. Exploratory Factor Analysis.....	28
C. Confirmatory Factor Analysis	30
D. Internal Consistency Results	31
E. Inter-Rater Agreement Results	32
F. Approach to Creating a Measure Score.....	35
G. Results of Sensitivity and Specificity Analyses	36
H. Stakeholder Feedback.....	37
I. Summary of Site Coordinator Debriefings.....	39
VI. CONCLUSIONS AND NEXT STEPS.....	41
REFERENCES.....	43
APPENDICES	
APPENDIX A. PTSD Tag1 Presentation	A-1
APPENDIX B. Technical Advisory Group Members and Affiliations	A-14
APPENDIX C. Technical Expert Panel Members.....	A-15

APPENDIX D. List of Reviewed PTSD Clinical Manuals	A-16
APPENDIX E. Participant Surveys.....	A-17
APPENDIX F. Data Processing and Flow Chart	A-33
APPENDIX G. Summary of Item-Level Responses	A-34
APPENDIX H. Exploratory and Confirmatory Factor Analysis Model Fit Statistics	A-42
APPENDIX I. Confirmatory Factor Analysis Model Solution	A-43
APPENDIX J. Summary of the Reliability Analyses.....	A-45
APPENDIX K. Item-Level Inter-Rater Agreement.....	A-49
APPENDIX L. Comparison Between Bayesian and Weighted Least Squares and Mean Variance Adjusted Estimators	A-56
APPENDIX M. ACI Index	A-60

LIST OF FIGURES AND TABLES

FIGURE IV.1. Clinician-Reported and Supervisor-Reported Educational Degree	17
FIGURE IV.2. Clinician-Reported and Supervisor-Reported Therapeutic Orientation.....	18
FIGURE IV.3. Sampling Process	20
FIGURE V.1. Distribution of Total Standardized Score by Respondent Type	35
FIGURE F.1. Data Processing and Flow Chart.....	A-33
TABLE II.1. Sources for Environmental Scan of PTSD Research Studies, Clinical Guidelines, and Quality Measures	3
TABLE II.2. Summary of PTSD TAG Members’ Prioritization of Measure Concepts	5
TABLE III.1. Examples of Clinician, Clinical Supervisor, and Client Survey Items	10
TABLE IV.1. Qualitative Research Questions	12
TABLE IV.2. Quantitative Pre-Testing and Analysis of Survey Measure	15
TABLE IV.3. Characteristics of Participating Clinicians and Supervisors by Site	16
TABLE IV.4. Summary of Completed Surveys	21
TABLE IV.5. Quantitative Pre-Testing and Analysis of Survey Measure	22
TABLE V.1. Summary of Survey Administration: Modes and Completion Times.....	28
TABLE V.2. The Five-Factor EFA Solution.....	30
TABLE V.3. Internal Consistency Results by Factor and Respondent	32
TABLE V.4. Inter-Rater Reliability for Clinicians, Supervisors, and Clients	32
TABLE V.5. Item-Level Inter-Rater Agreement Between Clinicians, Supervisors, and Clients	33
TABLE V.6. Results of Sensitivity and Specificity Analyses.....	37

TABLE B.1.	TAG Members and Affiliations	A-14
TABLE C.1.	TEP Members	A-15
TABLE G.1.	Summary of Item-Level Responses on the Clinician Survey	A-34
TABLE G.2.	Summary of Item-Level Responses on the Supervisor Survey.....	A-37
TABLE G.3.	Summary of Item-Level Responses on the Client Survey	A-40
TABLE H.1.	EFA Model Fit Statistics.....	A-42
TABLE H.2.	CFA Model Fit Statistics	A-42
TABLE I.1.	CFA Final Model Solution	A-43
TABLE J.1.	Summary of the Reliability Analysis for Factor 1: Structuring and Conducting the Session.....	A-45
TABLE J.2.	Summary of the Reliability Analysis for Factor 2: Psychoeducation and Teheapeutic Techniques.....	A-46
TABLE J.3.	Summary of the Reliability Analysis for Factor 3, Factor 4, and Factor 5	A-47
TABLE J.4.	Summary of Reliability Analysis by Measure Domain	A-48
TABLE K.1.	Item-Level Inter-Rator Agreement Between Supervisors and Clinicians.....	A-49
TABLE K.2.	Item-Level Inter-Rator Agreement Between Supervisors and Clients	A-52
TABLE K.3.	Item-Level Inter-Rator Agreement Between Clinicians and Clients	A-54
TABLE L.1.	Comparison Between Bayesian and WLSMV Adjusted Estimators: Clinician Confirmatory Factor-Analytic Models.....	A-56
TABLE L.2.	Comparison Between Bayesian and WLSMV Adjusted Estimators: Supervisor Confirmatory Factor-Analytic Models	A-58
TABLE L.3.	Comparison Between Bayesian and WLSMV Adjusted Estimators: Client Confirmatory Factor-Analytic Models.....	A-59

ACKNOWLEDGMENTS

Mathematica Policy Research and the National Committee for Quality Assurance prepared this report under contract to the Office of the Assistant Secretary for Planning and Evaluation (ASPE), U.S. Department of Health and Human Services (HHS) (HHSP23320100019WI/HHSP23337002T). Additional funding was provided by the HHS National Institute of Mental Health (NIMH). The authors appreciate the guidance of Kirsten Beronio, Joel Dubenitz, D.E.B. Potter (ASPE), and Joel Sherrill (NIMH). Jonathan Brown (Mathematica) provided feedback on this report and guidance throughout the project.

The views and opinions expressed here are those of the authors and do not necessarily reflect the views, opinions, or policies of ASPE, NIMH, or the technical expert panel. The authors are solely responsible for any errors.

ABSTRACT

Summary: Despite advances in the development of evidence-based treatment for adults with post-traumatic stress disorder (PTSD), the implementation of these treatments varies widely. To reduce this gap through wider dissemination of effective behavioral health treatment, the U.S. Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation and the HHS National Institute of Mental Health (NIMH) led a project that developed and pre-tested a quality measure of the delivery of psychotherapy for adults with PTSD that is concordant with evidence-based strategies.

Major findings: The project identified five measure constructs related to the delivery of evidence-based psychotherapy for PTSD: (1) structuring and conducting the therapy session; (2) psychoeducation and therapeutic techniques; (3) therapeutic alliance; (4) assessment; and (5) homework. The measure demonstrated fair to good reliability, but some items in the measure may be unnecessary or require refinement. Preliminary performance metrics were established that discriminate between clinicians who are high and low performers in the delivery of evidence-based psychotherapy. Stakeholders showed mixed support of the measure for quality improvement purposes; support for the measure's use in training and continuing education was strong.

Purpose: This project developed a measure of the delivery of evidence-based psychotherapy for adults with PTSD treated in ambulatory settings. The measure assesses care from three perspectives: the clinician, the clinical supervisor, and the client. The measure was pre-tested using quantitative and qualitative methods to assess attributes consistent with National Quality Forum endorsement criteria: importance, feasibility, usability, and scientific acceptability (reliability and validity).

Methods: This project first reviewed existing evidence and measures and gathered input from an advisory group to identify opportunities for new measures. Based on the evidence to support the measure concept, a survey was developed and pre-tested at six behavioral health organizations. Three parallel versions of the measure were developed and tested: clinician, supervisor, and client versions. Quantitative testing involved an examination of the measure's underlying constructs, estimation of its reliability, the creation of performance metrics, and calculation of the measure's sensitivity and specificity. Qualitative testing included focus groups with a range of stakeholders, as well as information gathering from test site coordinators, to obtain input on the measure's importance and face validity and to understand whether it could yield findings that could be used to inform quality improvement efforts.

ACRONYMS

The following acronyms are mentioned in this report and/or appendices.

ABBP	American Board of Behavioral Psychology
AC1	Adjusted for Chance-Corrected Statistic
AHRQ	HHS Agency for Healthcare Research and Quality
ASPE	HHS Office of the Assistant Secretary for Planning and Evaluation
BA	Bachelor's Degree
CASAC	Credentialed Alcoholism and Substance Abuse Counselor
CBT	Cognitive Behavioral Therapy
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CI	
CPT	Cognitive Processing Therapy
CQAIMH	Center for Quality Assessment in Mental Health
EFA	Exploratory Factor Analysis
EMDR	Eye Movement Desensitization and Reprocessing
HHS	U.S. Department of Health and Human Services
IRB	Institutional Review Board
KR20	Kuder-Richardson Formula 20
LCSW	Licensed Clinical Social Worker
LSW	Licensed Social Worker
NCQA	National Committee for Quality Assurance
NEIRB	New England Institutional Review Board
NIMH	HHS National Institute of Mental Health
NQF	National Quality Forum
NQMC	National Quality Measure Clearinghouse
OMB	Office of Management and Budget
P50	50th percentile
P75	75th percentile
PILOTS	Published International Literature on Traumatic Stress
PPP	Posterior Predictive P-value
PTSD	Post-Traumatic Stress Disorder
RMSEA	Root Mean Squared Error of Approximation

SAMHSA	HHS Substance Abuse and Mental Health Services Administration
SSRI	Selective Serotonin Reuptake Inhibitor
TAG	Technical Advisory Group
TEP	Technical Expert Panel
TLI	Tucker-Lewis Index
VA	U.S. Department of Veterans Affairs
WLSMV	Weighted Least Squares Means and Variance Adjusted estimator

EXECUTIVE SUMMARY

Purpose

In September 2011, the U.S. Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation, with support from the HHS National Institute of Mental Health, contracted with Mathematica Policy Research and the National Committee for Quality Assurance (NCQA) to develop quality measures for treatment of adults with PTSD. This 3.5-year project began by reviewing existing research evidence and measures and gathering input from a technical advisory group to identify and prioritize opportunities for new measures. We then specified and pre-tested a survey measure of the delivery of evidence-based psychotherapy for adults.

To develop the survey items, we sought input from a technical panel of experts in psychotherapeutic treatments for adults with PTSD and reviewed clinical manuals to produce a list of common evidence-based psychotherapeutic elements of PTSD. We converted the elements into three parallel sets of survey items to be completed by three different respondent groups: clinicians, clinical supervisors, and clients. The development of the three versions of the measure provides an opportunity to begin to assess which type(s) of rater results in the most credible and reliable measure. We revised the survey items based on input from groups of clinicians and clients. The clinician survey is presented in Appendix E.

To gather initial information about the measure's importance, feasibility, usability, and scientific acceptability in accordance with National Quality Forum endorsement standards, we gathered quantitative and qualitative data from six behavioral health organizations that provide outpatient services to adults with PTSD. Our quantitative testing involved fitting statistical models to identify the measure's underlying theoretical constructs and determine the necessity of each individual survey item. We examined the reliability of the measure using different psychometric tests depending on the type of reliability (inter-rater agreement or internal consistency) examined. We also conducted a preliminary assessment of the measure's sensitivity and specificity to determine the extent to which we could identify high-performing and low-performing clinicians, using scores we created based on performance at the 50th and 75th percentiles in the delivery of evidence-based psychotherapy. Finally, we conducted focus groups with a range of stakeholders and gathered information from site coordinators to obtain input on the measure's importance and face validity and to understand whether it could yield findings that could be used to inform quality improvement efforts. We also sought stakeholders' perspectives on practical barriers to implementing the measures.

Measure Testing Results

For each clinician, three therapy sessions for three different clients were sampled from the clinician's current caseload of adults with PTSD. The clinician, the clinician's supervisor, and the clients completed the survey following each sampled therapy session. We received 96

clinician, 97 supervisor, and 78 client surveys. Response rates were 98 percent, 99 percent, and 80 percent for clinicians, supervisors, and clients, respectively. The majority of clinicians and supervisors completed the survey on the web, whereas the majority of clients complete the survey on paper. On average, respondents completed the web survey in 8-10 minutes. In focus group discussions, most stakeholders felt the measure was too long and recommended shortening it.

We identified five similar underlying constructs in the measure that fit the data well in the clinician, supervisor, and client samples: (1) structuring and conducting the therapy session; (2) psychoeducation and therapeutic techniques; (3) therapeutic alliance; (4) assessment; and (5) homework. Some items correlated with more than one construct and other items had low correlations with the constructs. Taken together, the results suggest that the survey items assess constructs related to the delivery of psychotherapy for PTSD, but that some of the items may be unnecessary or require refinement. Although many stakeholders agreed the measure captures elements of psychotherapy, some stakeholders felt it focused too strongly on cognitive behavioral approaches when other psychotherapies are also delivered to adults with PTSD.

Across the reliability tests conducted, the measure demonstrated fair to good reliability. On average, we observed the highest reliability across all constructs in the supervisor sample, followed by the clinician and client samples. Supervisors and clinicians had the highest inter-rater agreement; supervisors and clients and clinicians and clients had comparable inter-rater agreement. The reliability results suggest some items may need revision, particularly among the items that comprise the “assessment” construct.

To begin to understand the measure’s validity, we calculated its sensitivity and specificity for each of the five constructs and compared clinician and client scores to the supervisor scores, which for the purposes of these analyses, we treated as a gold standard. We examined the implications for the measure’s sensitivity and specificity using two thresholds, the 50th and 75th percentiles, to determine high and low delivery of evidence-based psychotherapy. Measure sensitivity and specificity at the 50th percentile ranged from 0.50 to 0.79 and 0.49 to 0.78, respectively. At the 75th percentile, sensitivity ranged from 0.22 to 0.57 and specificity ranged from 0.75 to 0.85. Based on these preliminary findings, the 50th percentile threshold appears to better discriminate high and low performance. We treated the supervisor survey as a gold standard; however, stakeholders uniformly indicated a lack of endorsement for it due to the changes in process and the resources that would be required to routinely collect data from supervisors for quality improvement purposes. Some stakeholders noted a preference for the client survey, whereas others indicated a preference for either the client or clinician survey.

Conclusions and Next Steps

The development of a measure of the delivery of evidence-based psychotherapy has the potential to improve the quality of care for adults with PTSD. We made promising strides in creating the foundation of such a measure; however, a significant amount of additional work is needed to develop a final measure that can be used for accountability purposes. Below, we provide overarching conclusions and recommended next steps.

- **Additional input.** Although there was support for use of the measure in training and education, support for using it for accountability purposes was limited. Additional input from a larger group of stakeholders regarding the measure's use for internal quality improvement and the circumstances under which it would be useful would inform the next stages of measure development.
- **Further revisions.** Our analyses suggest that the survey assesses important underlying constructs associated with the delivery of evidence-based treatment for PTSD and that many survey items produce significant agreement across the three raters. The analyses also suggest that several items need refinement. For example, items with low inter-rater agreement and/or low internal consistencies may be candidates for deletion. Items with significant cross-loadings and moderate agreement could need revision. The surveys should be revised further, with additional cognitive testing and stakeholder input conducted on the refinements.
- **Further investigation of feasibility.** Several stakeholders expressed concern regarding the measure's feasibility. Refinements to the survey items may result in a shorter measure that takes less time to complete, which should improve the feasibility of using it. In addition, it would be useful to have additional information from a larger group of stakeholders regarding topics such as preferred survey mode (including mobile technology applications), the infrastructure available to support the measure, and approaches to automating aspects of site coordination.
- **Further development of the measure for broader application.** The factor analyses results identified therapeutic constructs that are likely relevant in the delivery of psychotherapy for conditions other than PTSD. The measure could be refined and further tested to create modules that broadly apply to the delivery of psychotherapy.
- **Examination of inter-rater reliability and factor structure using revised items and a larger sample.** Once the survey items have been refined, additional work will be needed to test whether the refinements improve inter-rater agreement and the factor structure. The goal of our current project was to pre-test this instrument. A pilot test with a larger sample that offers increased diversity in sites, clinicians, and clients would increase the external validity of the measure.
- **Examination of other scoring methods.** Our current thresholds for high and low delivery of evidence-based psychotherapy yielded positive results in terms of specificity and sensitivity. After item refinement, these scoring methods should be verified and compared to other possible methods of scoring. For example, contextual scoring may be beneficial, as it would allow clinicians flexibility in deviating from a treatment plan for appropriate reasons (such as in a case where a clinician did not use an expected set of therapeutic elements, because he or she had to help a client manage suicidal ideation).
- **Additional validity testing.** Additional psychometrics are needed to validate this measure. The use of an external, independent rater (not associated with the site) to serve as the preferred gold standard is important. To assess the measure's predictive validity,

information on patient outcomes (for example, symptom improvement, quality of life, and functioning) is critical.

The measure developed under this project has the potential to address significant gaps in quality of PTSD care. Additional work is needed to further prepare it for implementation on a larger-scale basis and to better understand the groups and situations where the measure will be most useful.

I. PROJECT RATIONALE

Post-traumatic stress disorder (PTSD) is a mental health disorder that sometimes results when individuals are directly or indirectly exposed to actual or threat of death, serious injury, or sexual violence (American Psychiatric Association 2013). An estimated 6.8 percent of the United States population has PTSD, with women estimated to have higher prevalence than men (9.7 percent for women versus 3.6 percent for men) (Kessler et al. 2005a; Kessler et al. 2005b) and veterans having a higher prevalence than the general population (7-20 percent for veterans of the recent wars, and estimates of about 30 percent for all veterans of the Vietnam War) (VA National Center for PTSD 2007, 2014).

Most people who experience traumatic events have a brief adjustment period during which they successfully cope with the experience. For others, symptoms worsen over time and last for months or years, disrupting their ability to function in everyday life. The cost of PTSD care can be significant. Studies have found that individuals with PTSD have increased health care service utilization, as measured by number of physical and mental health appointments and hospitalizations (Tuerk et al. 2012). The prevalence of PTSD among women with public insurance is over three times as high as for women with private insurance (Seng et al. 2009). Given the relatively higher risk of exposure to violence among people with low income, the need for effective PTSD treatment among Medicaid recipients is likely to be sizeable.

In recent years, increased national attention has led to an improvement in the types and effectiveness of treatments for individuals diagnosed with PTSD. Particularly promising are a number of psychotherapy treatment approaches -- for example, exposure therapy and cognitive processing therapy (CPT) -- that have demonstrated slightly to significantly better treatment outcome for those diagnosed with PTSD, such as reduction of symptoms and improved mental health.¹ Despite advances in the development of evidence-based treatment for adults with PTSD, the implementation of these treatments varies widely (Mellman et al. 2003), overall recovery rates remain low, and large disparities exist in the type and quality of mental health treatment across providers, patient populations, types of disorders, and even geographic regions. To enhance accountability, improve quality, and increase transparency for treatment of individuals with PTSD, the creation of quality measures is a first essential step. Well-constructed measures of evidence-based treatments could be used not only for overall quality improvement and monitoring purposes, but also for training and education and to determine the comparative effectiveness of treatments.

A. Project Purpose

In September 2011, the U.S. Department of Health and Human Services (HHS) Office of the Assistant Secretary for Planning and Evaluation (ASPE), with support from the HHS National Institute of Mental Health (NIMH), contracted with Mathematica Policy Research and

¹ For a recent synthesis of the evidence on PTSD treatment, see Institute of Medicine (2012, chapter 7).

the National Committee for Quality Assurance (NCQA) to develop quality measures for treatment of adults with PTSD. The Veterans Affairs and Military Health System have already invested significant resources to improve the care of active and retired duty individuals with PTSD; ASPE and NIMH were interested in building upon this existing work to develop measures that could be used in civilian ambulatory treatment settings. The overall goal was to develop measures that could eventually be used to hold providers or organizations accountable for delivering high quality care; however, there was recognition that PTSD quality of care measures could also be used for training and education and by other researchers.

The first step in this 3.5-year project involved prioritizing important measure concepts. Identification of measure gaps and priorities was informed through an environmental scan and input from a technical advisory group (TAG). The process identified several potential measure concepts, including measures that screen for common co-occurring conditions, assess appropriate receipt of psychotherapy and pharmacotherapy, routinely assess and monitor PTSD symptoms, and measure patient outcomes. The measure concept “the delivery of evidence-based psychotherapy” was selected. We then identified common elements of psychotherapy for PTSD with support from a newly formed technical expert panel (TEP),² developed measure specifications for a survey to assess the delivery of evidence-based psychotherapy for PTSD, and pre-tested the measure. The pre-testing involved quantitative data collection to examine the measure’s preliminary psychometric properties and explore potential approaches to scoring, as well as qualitative data collection, including focus groups and site coordinator debriefings to gather information on the measure’s feasibility, usefulness, and importance. Based on findings from the pre-testing, we recommended modifications to the measure specifications and additional testing of the measure to more fully understand its importance, scientific acceptability, usability, and feasibility as defined by the National Quality Forum (NQF).

B. Report Roadmap

This report summarizes the development and testing results of the quality measures for PTSD. Chapter II describes the process for selecting measure concepts. Chapter III explains the process for specifying the measures. Chapter IV describes the methods used to test the measure, and Chapters V summarizes the results. The final chapter offers conclusions and lessons learned from this project that may be applicable to future measure development and implementation efforts.

² TEP members were experts in psychotherapeutic treatments for adults with PTSD and were not members of the original TAG. They were selected to assist in identifying therapeutic elements and creating the initial measure.

II. SELECTION OF MEASURE CONCEPTS

The selection of measure concepts involved several steps: (1) conducting an environmental scan of evidence-based treatments for adults with PTSD; (2) conducting a review of existing measures of PTSD care to identify measurement gaps; and (3) convening a TAG to provide input on measure concepts and the evidence supporting those concepts. This chapter briefly describes these steps and how they influenced the development of the measure.

A. Environmental Scan of PTSD Treatments and Measures

After initial meetings with ASPE and NIMH to discuss priority measurement areas within the broad field of PTSD care and target populations for this quality measure development effort, we conducted a scan of research literature and clinical guidelines to identify evidence-based treatments for PTSD. The scan drew on systematic reviews (including meta-analyses), primary research studies, evidence-based clinical guidelines, and the recommendations of taskforces, including the Institute of Medicine’s taskforce on the treatment of PTSD (Institute of Medicine 2008, 2012).

Source of Information	Data Sources	Selected Search Terms
Research studies	<ul style="list-style-type: none"> • PubMed • Cochrane Database of Systematic Reviews • PsychINFO • National Center for the Study of PTSD website 	PTSD, trauma, psychotherapy, medication, drugs, pharmacotherapy, treatment, care, services
Clinical guidelines	<ul style="list-style-type: none"> • National Guideline Clearinghouse • Guidelines International Network • PubMed • Professional websites, for example, the American Psychiatric Association 	Psychology, psychiatry, adult and trauma, anxiety disorders, stress
Quality measures	<ul style="list-style-type: none"> • CQAIMH’s National Inventory of Mental Health Quality Measures database • NQF Quality Positioning System • PILOTS database • NQMC • Searches of behavioral health quality improvement initiatives, for example, the SAMHSA’s National Outcomes Measures • Conversations with PTSD experts 	Mental, behavioral, psychiatry, psychology, PTSD, trauma, anxiety, depression, substance, and patient experience, diabetes, cardiovascular

To identify relevant studies and guidelines, we developed search terms to guide this information gathering effort and identified data sources for the information (see Table II.1). We limited the scan to studies and guidelines in English and related to the treatment of PTSD in adults. We created detailed Excel spreadsheets with summaries of the treatment or intervention, the outcome measure(s), the results, and the study design and grading of the study design. We used this information to identify evidence-based treatments for adults with PTSD for which there was the strongest scientific evidence. Briefly, the results of the environmental scan identified strong evidence in support of the effectiveness of cognitive behavioral therapy (CBT), particularly exposure therapies, in the treatment of adults with PTSD. The scan also found clinical guideline support for -- but conflicting interpretations of -- the research on the effectiveness of selective serotonin reuptake inhibitors (SSRIs) and insufficient research evidence in adults with PTSD regarding the effectiveness of support services and care coordination (see Appendix A and the Institute of Medicine 2008, 2012).

B. Scan of Measures

We first began our search for quality measures of PTSD care similarly, by defining search terms (see Table II.1). We then searched the three most widely used sources of quality measures: the National Quality Measure Clearinghouse (NQMC), the NQF, and the online inventory maintained by the Center for Quality Assessment in Mental Health (CQAIMH). Additionally, we searched the Published International Literature on Traumatic Stress (PILOTS) database, which includes a large inventory of measures that are primarily used in research, and the HHS Substance Abuse and Mental Health Services Administration's (SAMHSA's) National Outcomes Measures. Our search for quality measures included measures related to PTSD care as well as ones related to physical or behavioral health conditions that commonly co-occur with PTSD (see Appendix A). We again summarized the information in an Excel spreadsheet that included information on the measure developers, specifications and data sources, NQF endorsement status, and level of evidence to support the measure.

C. Technical Advisory Group Review

The TAG was convened to provide input on the selection of measure concepts and available data sources to develop the measures. The group included research and clinical experts in the treatment of PTSD and behavioral health quality measurement. It also included a consumer representative as well as representatives from a health plan, the U.S. Department of Veterans Affairs (VA) health care system, and the community behavioral health system (see Appendix B for the list of TAG members).

The TAG meeting was held in March 2012. We summarized the evidence for PTSD care, and, based on that evidence, presented measure concepts for consideration in five broad domains: (1) psychotherapy; (2) pharmacotherapy; (3) assessment, monitoring, and treatment of commonly co-occurring behavioral and physical health conditions; (4) care coordination; and (5) consumer experiences with care. The TAG provided feedback on these measure concepts, suggested additional concepts, and offered input on the feasibility of developing the measures,

which rely on various data sources, including administrative data, electronic health records, medical record chart reviews, and survey data.

D. Selection of Measure Concept

To further refine the list of potential measure concepts for consideration, the TAG completed a measure prioritization exercise in mid-March where each member was independently asked to rate each concept on a 1-9 rating scale, with 1-3 classified as low priority, 4-6 as moderate priority, and 7-9 as high priority for each of the four NQF criteria (importance, scientific acceptability, usability, and feasibility; see Section IV).³ The TAG was asked to consider the availability of data, data collection burden, strength of the evidence supporting the concept, and saliency of the concept in prioritizing the concepts.

The TAG rated eight concepts as being of high importance; these included measures of psychotherapy, pharmacotherapy, screening for risk of suicide, and patient outcomes (Table II.2). Of these eight concepts, six were rated moderate feasibility and two (“receive at least eight sessions of CBT” and “receive CBT that includes specific components”) were rated low feasibility. None of the concepts was rated high feasibility. As noted in Table II.2, the TAG rated the other concepts to be of moderate importance.

TABLE II.2. Summary of PTSD TAG Members’ Prioritization of Measure Concepts					
Priority* Ranking	Concept	Importance Mean (range)	Scientific Acceptability Mean (range)	Usability Mean (range)	Feasibility Mean (range)
High priority					
1	Screened for risk of suicide	7.75 (6-9)	6.86 (4-8)	8.00 (7-9)	6.00 (5-8)
2	In psychotherapy and receive at least 8 sessions of CBT	7.50 (7-9)	5.00 (1-8)	8.14 (7-9)	3.71 (1-7)
3	In psychotherapy and receive CBT that includes specific components	7.38 (4-9)	5.57 (1-8)	7.29 (7-9)	2.57 (1-5)
4	Symptoms improve over a period of time	7.29 (4-9)	7.00 (4-9)	8.14 (4-9)	5.71 (4-8)
5	Receive CBT or SSRIs	7.25 (5-9)	5.29 (3-8)	8.00 (6-9)	5.29 (3-8)
6	Symptoms are assessed at routine intervals	7.13 (5-9)	7.29 (6-8)	7.43 (7-9)	5.57 (4-8)
7	On medication and receive SSRIs	7.00 (5-9)	6.71 (2-9)	7.29 (5-9)	7.29 (5-9)
8	On medication who receive a 4-month dosage of SSRIs	7.00 (4-9)	6.14 (2-9)	7.57 (6-9)	6.71 (4-9)
Moderate Priority					
9	In psychotherapy who receive CBT	6.88 (6-9)	4.00 (1-6)	7.29 (4-9)	3.57 (1-6)
10	Screened for depression	6.63 (4-9)	7.29 (6-8)	7.43 (5-9)	6.00 (4-7)
11	Functioning improves over a period of time	6.63 (4-9)	6.00 (2-8)	7.29 (5-9)	5.29 (3-7)
12	PTSD screened for substance abuse	6.50 (3-9)	6.86 (4-8)	7.86 (7-9)	5.00 (2-7)

³ The 1-9 rating scale follows the ratings practices used by RAND in other similar prioritization and appropriateness ratings exercises (AHRQ n.d.; Brook et al. 1990; Fitch et al. 2001).

TABLE II.2 (continued)					
Priority* Ranking	Concept	Importance Mean (range)	Scientific Acceptability Mean (range)	Usability Mean (range)	Feasibility Mean (range)
13	Quality of life improves over a period of time	6.50 (4-9)	5.71 (2-8)	6.86 (5-9)	5.14 (3-7)
14	No improvement or a worsening of symptoms, and have a documented change in treatment approach	6.38 (2-9)	6.57 (3-8)	7.14 (4-9)	4.86 (4-7)
15	Assessed for sleep problems	6.38 (3-9)	6.86 (6-8)	7.57 (6-9)	5.00 (2-7)
16	Quality of life and functioning are assessed at routine intervals	6.38 (2-9)	5.71 (2-7)	7.14 (6-9)	4.71 (2-7)
17	Treatment options such as psychotherapy, medications, or a combination discussed	6.25 (3-9)	5.71 (3-7)	6.86 (5-9)	4.14 (2-7)
18	On medication and assessed regularly for medication side effects	6.13 (4-9)	6.57 (5-8)	7.14 (5-9)	3.71 (2-6)
19	Adults with documented comorbidities who have a documented care management/ coordination plan	6.00 (4-9)	5.29 (4-7)	6.57 (3-9)	4.29 (2-7)
20	Treatment preferences were considered	5.88 (3-9)	5.00 (3-7)	6.86 (5-9)	4.00 (2-6)
21	Needs for support services have been assessed	5.88 (2-9)	5.29 (2-7)	6.71 (4-9)	4.00 (2-7)
22	On medication who have a documented assessment of medication possession ratio (or other measure of medication adherence)	5.75 (1-9)	5.50 (3-7)	7.17 (6-9)	5.67 (2-8)
23	On multiple medications who have documentation of an assessment for potential drug interactions	5.63 (3-9)	6.57 (5-9)	6.57 (5-9)	4.14 (2-7)
24	Receive a referral and have documentation that the referral was followed up	5.63 (2-9)	5.57 (3-8)	6.71 (4-9)	4.43 (1-7)
25	Screened for pain	5.63 (3-9)	6.00 (2-7)	6.14 (2-9)	5.14 (2-7)
26	Receive care from more than 1 provider--that has been communicated to all providers	5.50 (2-9)	5.71 (2-7)	6.43 (3-9)	4.00 (1-7)
27	Assessed to determine if care management/care coordination is needed	5.50 (2-9)	5.50 (4-7)	6.33 (3-9)	4.00 (2-7)
28	Screened for glucose levels, lipids, high blood pressure	5.25 (3-7)	6.17 (2-9)	5.67 (2-8)	6.67 (5-8)
29	Receive antipsychotic medication	3.75 (2-8)	5.86 (2-8)	5.29 (3-8)	6.86 (4-8)

* Based upon the NQF "importance" criteria.

The TAG’s identification of the eight priority measure concepts provided valuable information to inform a discussion between Mathematica, ASPE, and NIMH regarding the subsequent direction and focus of the project. Together, we selected the “delivery of evidence-based psychotherapy” concept for development and specification. This decision was influenced by the strength of the evidence regarding CBTs as the recommended first line of treatment for adults with PTSD, as well as the limited national data on the quality of psychotherapy treatment.

This gap provides ASPE and NIMH the opportunity to not only advance the state of knowledge regarding the quality of psychotherapy delivered to individuals with PTSD, but also inform the broader mental health and quality improvement fields in approaches to measuring quality of psychotherapy for other mental disorders.

III. MEASURE SPECIFICATION

A. Selection of Data Source

Based on feedback from the TAG and clinical and quality measure development experts, we determined that the information needed to calculate the measure was not available from claims or medical records and would therefore require primary data collection in the form of a survey. To reduce respondent burden, and based upon feedback from clinicians with data collection expertise, we also determined the survey would be web-based, with paper versions provided upon request. Below, we summarize the other data sources considered.

- **Administrative claims data.** Although administrative claims-based measures require comparatively lower levels of resources from organizations than measures that utilize other data sources, there are no data on the specific elements of psychotherapy captured in claims.
- **Health record data.** Information on the use of specific psychotherapies is sometimes documented in clinician case notes; however, there is a lack of standardization in the type and specificity of information provided. Further, TAG members expressed concerns that the cost associated with reliably manually abstracting the necessary information would likely present significant barriers to the adoption of the measure. The use of electronic health records to capture information on the use of specific psychotherapies may increase the feasibility and reliability of the measure; however, the current state of electronic health records in the mental health field does not support the implementation of this type of measure at this time.

Although implementing surveys can be resource intensive relative to measures developed with other data sources, they provide a forum to collect treatment implementation information that is not available in administrative or health record data. They also have the added benefit of providing a mechanism to gather information on the quality of psychotherapy from multiple stakeholder perspectives.

B. Identification of Critical PTSD Psychotherapy Treatment Elements using an Established Methodology

Convening a TEP to generate a list of common treatment elements. As a first step in the development of this survey measure, we used the established “distillation and matching” approach (Chorpita 2005, 2009) to identify the elements present in evidence-based psychotherapy for adults with PTSD. Given the current research evidence, we focused on elements of cognitive behavioral approaches to the treatment of PTSD. In accordance with this method, we convened a (new) TEP, composed of national and international experts in the treatment of PTSD, particularly in prolonged exposure therapy and CPT, two psychotherapies

that fall under the broader umbrella of cognitive behavioral approaches (see Appendix C). The TEP recommended an initial list of psychotherapy treatment elements that largely draw from these two therapies and include elements such as the use of Socratic questioning, cognitive restructuring, and homework assignments.

PTSD clinical treatment manual review. To determine the extent to which the psychotherapy treatment elements commonly occur, we systematically reviewed PTSD clinical manuals for the presence of the identified elements. We identified eight PTSD treatment clinical manuals (Appendix D) through web-based searches and recommendations from PTSD clinical experts. Two independent reviewers from Mathematica read each manual and documented the presence or absence of each treatment element. Mathematica’s project director or deputy project director resolved discrepancies between the reviewers. In total, reviewers identified 30 elements, agreeing upon their presence or absence for 23 of them (77 percent).

Selection of common psychotherapy elements. To identify the final list of common evidence-based psychotherapy elements for the treatment of PTSD, we reviewed the frequency with which the elements were identified in the clinical manuals. We dropped any treatment elements such as stress inoculation training and relaxation training that were identified in three or fewer clinical manuals, and did not translate them into survey items. We then convened the TEP to provide input on the final list. They generally agreed with our identification and prioritization of treatment elements and recommended the inclusion of two additional elements: assessing and monitoring client symptoms and being directive in therapy sessions, which we incorporated into the list. The TEP also provided input on the extent to which the identified elements assess underlying constructs of evidence-based psychotherapy. Although they agreed that the identified elements could be grouped into treatment constructs, they were unable to reach consensus regarding how they should be grouped. The final list included 35 common elements.

C. Survey Item Development

Once we finalized the common elements, we converted them into survey items (see Appendix E). Three items drawn from another instrument were measured on a Likert scale; the remaining items consisted of categorical “yes or no” response options, with options for “don’t know” or “don’t remember.” For example, to assess the element “challenging the clients’ problematic beliefs using the Socratic method,” we developed the question:

Did you use a Socratic discussion method, that is, statements or questions designed for the client to examine their beliefs?

For example:

- How do you know this? Can you give me an example?
- What are some other ways of viewing this? What are the pros and cons to your way of thinking about this?
- How did you come to this conclusion? What evidence do you have to justify this?

We developed three parallel versions of the items to be completed by three different respondents within 24 hours of a sampled therapy session: clinicians, clinical supervisors, and clients (see Chapter IV for more information on the sampling design). The development of the three versions of the measure provides an opportunity to begin to assess which type of rater(s) results in the most credible and reliable measure. Table III.1 illustrates how the wording of the items differs based upon the rater.

TABLE III.1. Examples of Clinician, Clinical Supervisor, and Client Survey Items			
Common Element	Clinician	Clinical Supervisor	Client
Agenda setting	Did you set an agenda?	Did the therapist set an agenda?	Did you and your therapist discuss an agenda or plan for your session?
Socratic questioning	Did you use a Socratic discussion method, that is, statements or questions designed for the client to examine his/her beliefs?	Did the therapist use a Socratic discussion method, that is, statements or questions designed for the client to examine his/her beliefs?	Did your therapist ask you several direct questions to make you think critically about or examine your thoughts, feelings, or beliefs?
Risk assessment	Did you conduct a suicide risk assessment for this client?	Did the therapist conduct suicide risk assessment during this session?	During this session, did your therapist ask you if you had thoughts about committing suicide?

In addition to modifications to the wording of the items, the client version also had a reduced number of items. Based upon recommendations from the TEP, we removed items that the TEP believed clients would have difficulty addressing (for example, the use of cognitive restructuring techniques) and combined related items (for example, setting an agenda and reviewing the agenda). The resulting client version included 25 items.

IV. APPROACH TO MEASURE TESTING

Following the specification of the measure, we pre-tested it in two stages. In the first stage, we used qualitative methods to gather information on the importance and usefulness of the measure, the validity of the survey items, and the feasibility of using this type of measure.

In the second stage, we used both quantitative and qualitative methods. Stage 2 of the pre-testing effort was designed to gather information to inform potential uses of the measure, assess preliminary information on its psychometric properties, explore approaches to developing a measure score, and gather information on the measure's implementation.

Although it is formative in nature, our testing initiative was designed to lay a foundation for additional measure testing, and if, appropriate, for future endorsement by the NQF. As such, we framed many of the research questions around the NQF measure criteria *importance*, *scientific acceptability*, *usability*, and *feasibility*, defined by the NQF as:

- **Importance.** The strength of evidence supporting that a measure concept promotes high quality care and allows for differentiation in performance.
- **Scientific acceptability.** The verification that the psychometric properties of a measure--validity and reliability--are strong enough to justify its use to assess quality of care:
 - **Validity.** The ability of measure specifications to promote *accuracy* in data collection and measure score calculation to ensure appropriate characterization of performance.
 - **Reliability.** The ability of measure specifications to promote *consistency* in data collection and aggregation to ensure that variability in measure score reflects actual variation in performance.
- **Usability.** The value of a measure in informing quality improvement activities.
- **Feasibility.** The availability of data elements required for the calculation of a measure, whether a measure is susceptible to inaccuracies, and the level of effort involved in collecting and calculating the measure.

This chapter describes the methods used to test each of these criteria. We briefly summarize the testing questions and specific methods for each stage of measure testing.

A. Stage 1--Testing the Survey Items

After we developed the survey items, we then gathered qualitative information to answer the research questions in Table IV.1. At this stage, the priority was to gather information regarding the face validity and interpretability of the survey items so we could refine them, as needed, in preparation for more formal measure testing in Stage 2. We also gathered preliminary

stakeholder input on the measure’s importance, usability, and feasibility. Below, we describe the process we used to gather qualitative information, and then summarize the information learned.

TABLE IV.1. Qualitative Research Questions	
Criterion	Testing Question
Importance	Is the measure appropriate for assessing quality of care?
Validity	Are stakeholders interpreting the survey items in the way we intended them to be interpreted? Are there alternate ways of wording key concepts that better resonate with stakeholders? Do the survey items measure quality of psychotherapy?
Usability	How would different organization and entities use this measure to improve the quality of PTSD care?
Feasibility	How burdensome is the measure to complete? Can the measure be accurately scored?

1. Telephone Discussions

From May 2013 to October 2013, we hosted a series of telephone discussions with stakeholders to gather input on the wording and interpretation of the survey items, the usefulness of the measure to improving quality of PTSD care, and the feasibility of completing the measure. We utilized an iterative process to gathering information whereby we held a discussion with stakeholders, revised the survey items based upon the feedback received, and then held additional discussions with new stakeholders. Discussion group participants represented two types of stakeholder groups:

- **Clinicians and clinical supervisors.** Participants included nine clinicians and clinical supervisors who were experienced in providing psychotherapy to adults with PTSD. They were identified through recommendations by project team members and members of the TEP. We hosted a total of four discussions with clinicians and supervisors, with group sizes ranging from one to three participants.
- **Clients.** Participants included six adults (three men and three women) who were either nearing the end of treatment for PTSD or had completed treatment within the previous year. They included both military veterans and civilians and were identified through recommendations from TEP members, ASPE, project team members, and postings on listservs such as the National Alliance for the Mentally Ill’s listserv. We hosted three discussions with clients, with group sizes ranging from one to two participants. Participants received a \$20 gift card.

We drafted specific questions to fit the particular expertise of each type of discussion group and revised the questions for each subsequent discussion.

2. Summary of Stakeholder Input

Survey items. Participants in both stakeholder groups provided valuable suggestions on how to improve the clarity and meaning of the survey items. Based upon their input, we altered the wording of some items, added concrete examples to further clarify the items, and combined redundant items.

Importance and usefulness. There was general agreement on the importance of improving the quality of PTSD care; however, not all participants saw value in this specific measure. Some clinicians argued that the measure only assesses the delivery of CBT when other therapies are also effective in treating PTSD; as a result, the measure would not be useful or of interest to clinicians who do not provide CBT. Others, particularly clinical supervisors, felt it would be useful. Some clients indicated the importance of improved outcomes rather than the therapeutic techniques used by their clinicians. Other clients appreciated the effort to improve the quality of their care and saw value in the measure.

Feasibility. Stakeholder feedback on the feasibility of using the measure varied. Some clinicians and supervisors expressed concern regarding the length of the measure, the time required to complete it, and the feasibility of completing it within 24 hours of a given therapy session. Clients did not indicate concerns with the measure's length.

3. Development of the Final Survey

Once we completed revisions to the survey items based on feedback obtained from the telephone discussions, we provided TEP members with the opportunity to review and comment on a revised version of the survey. ASPE and NIMH also conducted a final review. We then made minor revisions and finalized the survey items, which consisted of 32 items in the clinician and supervisor versions and 25 items in the client version. The final surveys are available in Appendix E.

Once the survey items were finalized, we created web-based surveys using *Opinio* software (Version 6.7.1; ObjectPlanet, Inc., Norway). Prior to launching data collection, we rigorously tested the program to ensure that response fields functioned properly; users could move back and forth between questions, change answers, and save and return to the survey to complete at a later time; and entered responses were correctly coded and stored.

B. Stage 2--Pre-Testing the Measure

Once we finalized the development of the surveys, we collected quantitative and qualitative data to pre-test the measure. The quantitative data collection involved administering the surveys at specialty behavioral health organizations to assess the psychometric properties of the measure, potential approaches to scoring the measure, and potential implementation challenges. The qualitative data collection involved gathering feedback from stakeholder focus groups and individuals who coordinated measure testing within their organization to assess the measure's usefulness and feasibility.

We first describe our approach to quantitative testing and then our approach to qualitative testing.

1. Quantitative Testing of the Survey Measure

The quantitative testing was designed to answer the questions in Table IV.2. We pre-tested the measure at six behavioral health organizations, which allowed us to assess the organizations' abilities to collect the data, the initial psychometric properties of the measure, and different strategies for calculating a measure score. There are three key features of the quantitative testing design:

- **Survey completion by multiple respondent types.** There is a dearth of empirical evidence to suggest which type of respondent will produce the most credible and reliable information on the delivery of evidence-based psychotherapy. Some stakeholders who participated in Stage 1 of measure testing, as well as some TAG and TEP members, suggest that clinicians may over-report the delivery of evidence-based therapeutic elements. Others suggested that clients may have difficulty in recognizing technical aspects of the therapeutic elements while they are in the midst of therapy, and may under-report the delivery of evidence-based therapeutic elements. To inform future decisions regarding the optimal respondent type, clinicians, their supervisors, and a sample of their clients completed the survey on the same sampled therapy sessions (see Section IV.B.4 for information on the sampling design). For the purposes of this data collection effort, we considered supervisors to be the most experienced and objective raters and treated their responses as the gold standard. As such, clinician and client responses were compared to supervisor responses.
- **Survey completion at multiple stages of treatment.** Cognitive behavioral approaches to treating PTSD typically follow a general sequence of events. There may be appropriate variation in when specific therapeutic elements are delivered; however, one might expect certain items to be delivered rarely at, for example, the beginning or end of treatment. To begin to develop a rich understanding of the delivery of evidence-based psychotherapy across the course of treatment, clinicians and their supervisors completed the survey following three therapy sessions of clients who were at different stages in the therapy process -- beginning, middle, and end. Clients completed the survey only once.
- **Survey completion by clinicians and supervisors who represent a range of therapeutic orientations.** Although the majority of the survey items reflect cognitive behavioral approaches, we recruited organizations that employ clinicians who utilized CBT as well as other types of psychotherapy in the treatment of individuals with PTSD. Obtaining this range of techniques was necessary to assess how the measure performs.

TABLE IV.2. Quantitative Pre-Testing and Analysis of Survey Measure		
Letter Name	Uppercase	Lowercase
Importance	Does performance on the measure vary? How does performance vary when different approaches to scoring the measure are applied?	Descriptive analysis (mean, range, outliers) of performance
Factor-analytic structure	How many underlying psychotherapeutic constructs does the measure include? What does the factor structure imply regarding the number of items in measure?	EFA and CFA
Reliability: Internal consistency	What is the extent of the agreement between the items in each identified factor?	Alpha statistic
Reliability: Inter-rater	To what extent is there agreement between clinicians, supervisors, and clients in rating the survey items and in the overall survey?	Agreement using AC1 statistic
Validity	To what extent does the measure distinguish between clinicians who do and do not deliver evidence-based psychotherapy?	Sensitivity and specificity analyses
Feasibility	On average, how long did it take participants to complete the measure?	Descriptive analysis (means and ranges)

Here we describe the characteristics of the participating behavioral health organizations and data collection process.

2. Site Characteristics

From June 2014 to January 2015, we sought to recruit 36 clinicians employed by behavioral health organizations that delivered psychotherapy to adults with PTSD in outpatient treatment settings. We announced the project via the listservs of the National Council on Community Behavioral Health, American Counseling Association, and Kent State Counselor Education and Supervision. We also contacted organizations recommended by members of the TEP and project team. We identified other potential organizations through web-based searches.

As organizations expressed interest, we conducted informational meetings where we provided additional information regarding the project and its goals, and specifics about the testing activities. We then assessed whether the interested organizations met the desired requirements, employing the following:

- Clinicians who provide psychotherapy to at least three adult clients (in various phases of treatment) with a diagnosis of PTSD.
- Clinical supervisors who routinely provide clinical supervision via direct observation or video or audio tape, or were willing to provide these types of supervision for selected therapy sessions.

- An individual within the organization able and willing to coordinate data collection activities for their organization, including client recruitment.

We conducted follow-up interviews to gather additional information on the number of eligible clinicians and supervisors, the type of psychotherapy provided to adults with PTSD, and the type and frequency of supervision. We confirmed that they had the capacity to participate in the testing and discussed potential challenges to their participation before selecting the final organizations. We then established a Memorandum of Understanding and a Business Associate Agreement with each organization to govern the secure use of the data submitted under this project. We provided each organization with a modest honorarium to offset the costs of data collection. Where necessary, we submitted Institutional Review Board (IRB) materials for review by organizations' internal IRBs.

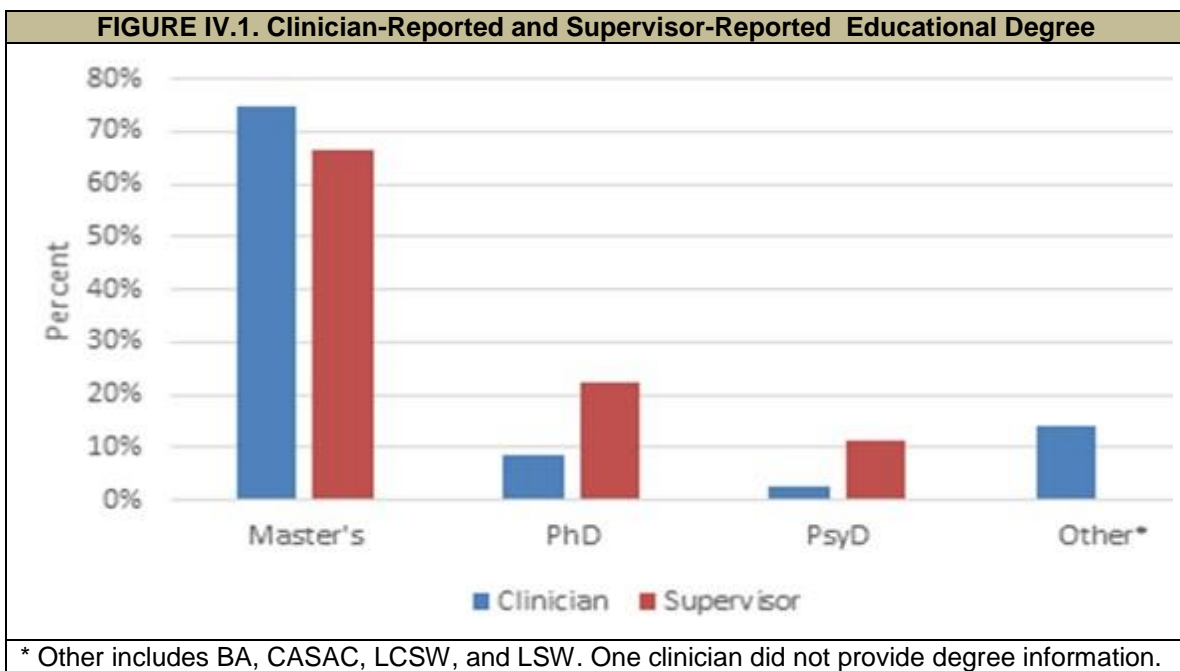
In total, we recruited six behavioral health organizations with a total of 37 clinicians and nine clinical supervisors. The behavioral health organizations were located in the Midwest and on the East Coast; most served individuals with public and private insurance.

3. *Clinician and Supervisor Characteristics*

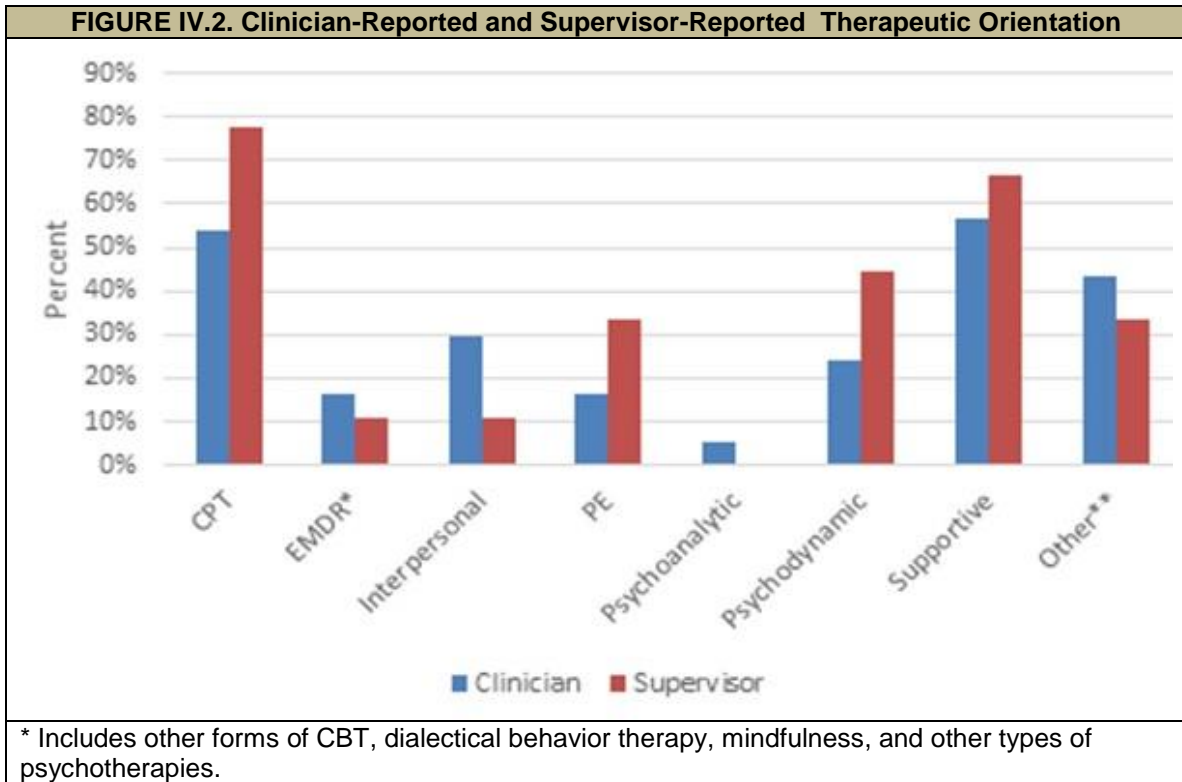
TABLE IV.3. Characteristics of Participating Clinicians and Supervisors by Site								
		Sample Size	Average Number of Years Providing Therapy (range)	Average Number of Years Providing Treatment for PTSD (range)	Average Current Number of Clients Per Clinician (range)	Current Number of Clients with PTSD (range)	Percentage Currently Licensed	Percentage with Accreditations or Certifications in CBT
Total	Clinicians	37	7.5 (1-29)	6.4 (0-29)	50 (7-100)	11 (0-40)	70.3	67.6%
	Supervisors	9	16 (4-30)	10.7 (2-26)	20 (0-40)	4 (0-15)	100%	88.9%
Site A	Clinicians	11	2.6 (1-7)	2.6 (1-7)	33 (20-45)	6 (0-10)	63.6%	54.5%
	Supervisors	2	8 (4-12)	3 (2-4)	27 (25-28)	3 (2-4)	100%	100%
Site B	Clinicians	3	3 (1-5)	5 (5-5)	25 (7-60)	2 (0-3)	66.7%	66.7%
	Supervisors	2	8 (6-10)	4.5 (4-5)	7 (6-8)	4 (3-5)	100%	100%
Site C	Clinicians	5	6.6 (5-8)	2.4 (1-4)	24 (12-29)	24 (12-29)	100%	80%
	Supervisors	1	14 (n=1)	4 (n=1)	17 (n=1)	15 (n=1)	100%	100%
Site D	Clinicians	6	12.7 (2-29)	10.2 (2-29)	58 (40-75)	9 (3-20)	66.7%	100%
	Supervisors	1	18 (n=1)	18 (n=1)	18 (n=1)	3 (n=1)	100%	100%
Site E	Clinicians	7	9.3 (1-20)	9.14 (0-20)	100 (99-100)	19 (5-40)	71.4%	57.1%
	Supervisors	1	25 (n=1)	25 (n=1)	0 (n=1)	0 (n=1)	100%	100%
Site F	Clinicians	5	13.2 (5-23)	12.2 (3-21)	53 (35-70)	6 (3-9)	60%	60%
	Supervisors	2	27.5 (25-30)	17 (8-26)	40 (40-40)	4 (4-4)	100%	50%

As described in Table IV.3, the clinicians who completed the survey, on average, had been providing therapy for 7.5 years and providing treatment for PTSD for 6.4 years. Clinicians' current caseloads averaged 50 clients per clinician; almost 25 percent of those clients had PTSD. On average, participating supervisors had been providing therapy for 16 years and providing treatment for PTSD for 10.7 years. Supervisors saw an average of 20 clients, including an average of four clients with PTSD. The majority of participating clinicians (70.3 percent) and all of the supervisors (100 percent) were currently licensed as mental health professionals. The majority of clinicians and supervisors were also accredited or certified in cognitive behavior therapy (67.6 percent of supervisors and 88.9 percent of clinicians).

The most common degree type was a master's degree, held by 75 percent and 67 percent of clinicians and supervisors, respectively (see Figure IV.1).



Over half (54 percent) of the clinicians identified their therapeutic orientation as “supportive,” whereas the majority of supervisors (78 percent) identified CPT, a form of CBT, as their therapeutic orientation (see Figure IV.2).



4. Data Collection Process

Site coordinator training. To facilitate the data collection process, we asked each participating organization to identify a staff member to serve as a site coordinator. These individuals filled a critical role. Their responsibilities included providing Mathematica with the information on eligible clinicians, supervisors, and clients to draw a study sample; notifying clinicians and supervisors when they were due to complete a survey; providing follow-up reminders to clinicians and supervisors to complete past-due surveys; describing the project and data collection effort to eligible clients; and attending regular meetings with Mathematica/NCQA.

To prepare the site coordinators' for their involvement in the project, we held web-based trainings. In these trainings, we oriented the coordinators to the goals and objectives of the project and their role and responsibilities on the project. We provided guidelines and tips for communicating with clinicians, supervisors, and clients, instruction on how to access the survey, and best practices for data security. We also provided them with a packet of materials to facilitate completion of their tasks.

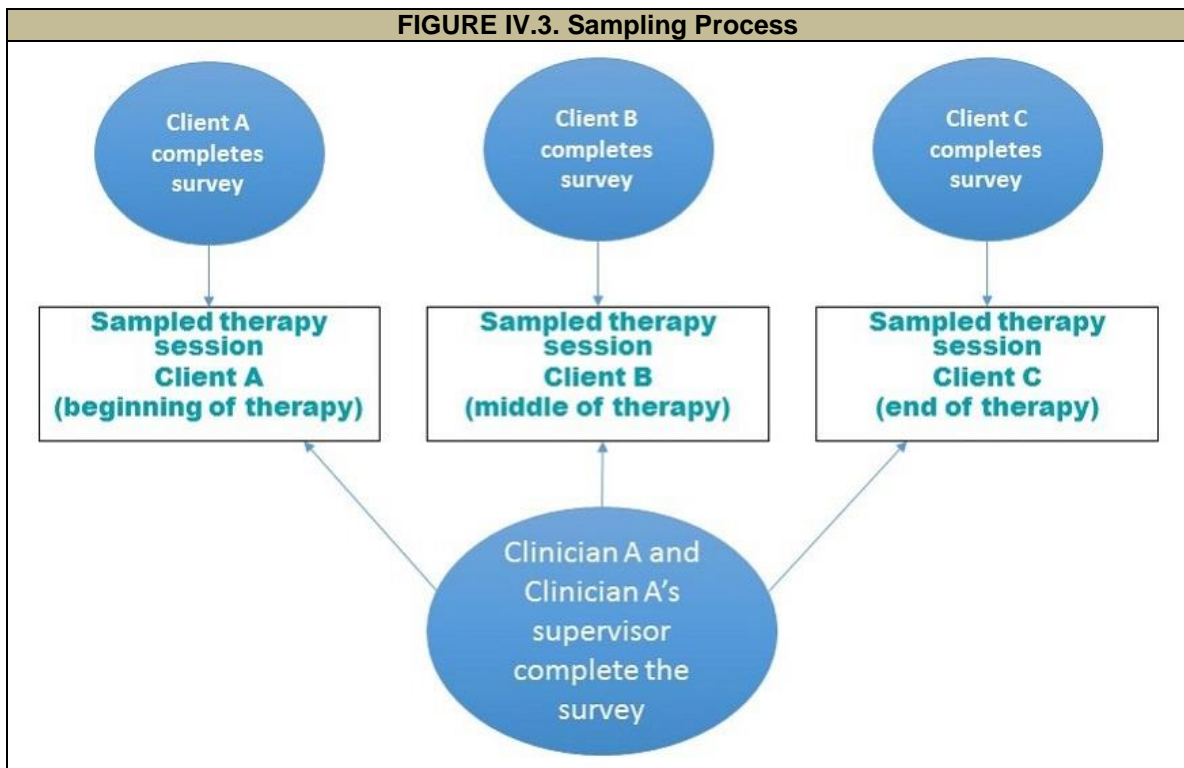
To further support the site coordinators, Mathematica/NCQA held frequent communication with them. Project staff emailed site coordinators no less frequently than every other day to provide updates on each site's response rates, confirm upcoming therapy session dates, and, if needed, determine if resampling was necessary due to missed therapy appointments or a client terminating therapy. They also held weekly group meetings with the sites to discuss the status of data collection activities and to collectively strategize approaches for collecting data.

Sample selection and survey administration. To select the study sample, site coordinators securely transmitted to Mathematica a list of clinicians who were currently providing psychotherapy to at least three adults with PTSD, their supervisors, and their clients. The site coordinators also provided information on the clients' treatment start date, expected length of treatment, and date of next therapy session. Mathematica, with input from the site coordinators, then classified the clients' upcoming therapy session as occurring in the beginning, middle, or end of treatment, and drew a study sample following the process described below and illustrated in Figure IV.3:

- For each clinician, three therapy sessions were sampled from the clinician's current caseload of adults with PTSD -- one therapy session of a client who recently started therapy, a second therapy session of another client who was in the middle of therapy, and a third therapy session of another client who was toward the end of therapy.
 - The clinician completed the survey following each of the three sampled therapy sessions.
 - Clinicians were instructed to complete the survey within 24 hours of each sampled therapy session.
- The clinician's supervisor was also sampled and also completed the survey following each of the three sampled therapy sessions.
 - Most of the participating supervisors supervised more than one participating clinician. The number of surveys a supervisor completed therefore depended on the number of participating clinicians he or she supervised. For example, a supervisor who supervised one clinician completed the survey three times, whereas, a supervisor who supervised three clinicians completed the survey nine times (three times on three different therapy sessions per clinician).
 - Supervisors were instructed to complete the survey within 24 hours of audio taper review or direction observation of the sampled therapy session.
- The clients attending each of the sampled therapy sessions were also sampled. They completed the survey once, following the sampled therapy session.
 - If the client refused to participate in the project, the sampled therapy session was discarded; neither the clinician nor his or her supervisor completed the survey on the session. Instead, we resampled a therapy session from another client on the same clinician's caseload, if possible. In nine cases, the clinicians did not have another client in the appropriate stage of treatment to resample.
 - If the client discontinued treatment or missed three consecutive appointments, which the site coordinators suggested was an indication of passively discontinuing treatment, the therapy session was discarded. A therapy session from another client on the clinician's caseload was sampled, if possible. In 16 cases, clinicians did not have another client in the appropriate stage of treatment to resample.
- The sampling structure resulted in survey responses on the same therapy session from clinicians, supervisors, and clients.

Once the therapy sessions were sampled, Mathematica sent each site coordinator a file with the names of the selected clients and therapy session dates, as well as direct web survey links for use by the clinicians, supervisors, and clients. Site coordinators then distributed paper and/or electronic survey alerts to participating staff 48 hours before and on the day of a selected session to remind them of the need to complete the survey following the selected therapy session. Site coordinators provided follow-up reminder letters and/or emails to staff with delayed survey responses. Appendix F depicts the data collection process.

When sampled clients checked in for their appointment, site coordinators described the project and its associated risks and benefits and invited them to participate. Clients were provided with written information about the project, information on how to access the survey online, and, if desired, a paper copy of the survey with a pre-paid return addressed envelope. In sites with local computers, clients were also given the option to complete the survey on-site prior to leaving.



Summary of response rates by site. A total of 144 therapy sessions were sampled (see Table IV.4). After accounting for attrition and refusals to participate in the project, 98 percent of clinicians, 99 percent of supervisors, and 80 percent of clients completed the survey. One clinician and one supervisor dropped from the study; new or already participating staff replaced them. Over 25 percent of sampled clients discontinued treatment or missed three consecutive therapy sessions; however, in over half of those cases, we were able to sample a replacement client from the clinician’s caseload.

TABLE IV.4. Summary of Completed Surveys									
		Total Number of Sampled Sessions	Attrition from Treatment with Replacement*	Attrition from Treatment without Replacement	Clients Declined to Participate with Replacement	Clients Declined to Participate without Replacement	Total Expected Completed Surveys	Number of Completed Surveys	Response Rate
Total	Clinicians	144	1	0	NA	NA	98	96	98%
	Supervisors	144	1	0	NA	NA	98	97	99%
	Clients	144	21	15	0	11	97**	78	80%
Site A	Clinicians	42	0	0	NA	NA	34	34	100%
	Supervisors	42	0	0	NA	NA	34	34	100%
	Clients	42	6	1	0	2	34	23	68%
Site B	Clinicians	18	0	0	NA	NA	10	10	100%
	Supervisors	18	0	0	NA	NA	10	10	100%
	Clients	18	3	4	0	1	10	8	80%
Site C	Clinicians	22	0	0	NA	NA	15	15	100%
	Supervisors	22	0	0	NA	NA	15	15	100%
	Clients	22	7	0	0	0	15	14	93%
Site D	Clinicians	21	0	0	NA	NA	14	12	86%
	Supervisors	21	1	0	NA	NA	14	14	100%
	Clients	21	0	4	0	3	14	13	98%
Site E	Clinicians	23	0	0	NA	NA	18	18	100%
	Supervisors	23	0	0	NA	NA	18	17	94%
	Clients	23	2	3	0	0	18	17	94%
Site F	Clinicians	18	1	0	NA	NA	7	7	100%
	Supervisors	18	0	0	NA	NA	7	7	100%
	Clients	18	3	3	0	5	7	4	57%

* Attribution is defined as discontinuing treatment or more than 3 consecutive missed therapy sessions.
** Note that 1 participant's refusal was mailed in after the clinician and supervisor had completed their surveys.

5. Quantitative Analysis

The quantitative analyses were designed to answer the questions in Table IV.5.

a. Quantitative testing of the measure's theoretical structure

To identify the measure's theoretical structure and assess the necessity of each survey item across clinicians, supervisors, and clients, we conducted an exploratory factor analysis (EFA) and then used the resulting EFA model as a basis for confirmatory factor analyses (CFA).

Exploratory factor analysis. Factor analysis is a data-reduction tool commonly used in measure development. It is used to examine the variability and correlation among survey items to determine if a smaller pool of items (or factors) is being measured by the items. EFA is a data-driven approach that imposes no restrictions on the data, such as pre-existing ideas about the number of constructs in the measure or the patterns of relationships between the survey items. To identify the measure's underlying structure in the EFA, we combined the clinician, supervisor, and client survey item responses. In this stage, we did not account for respondent type, but rather wanted to examine the overall factor structure. In CFA (described below), we conducted separate analyses by respondent type. We used the default oblique Geomin factor rotation method. This rotation method assumes correlation between factors but is equally robust if the factors are not sufficiently correlated or not correlated at all. Because the factor-analytic model included categorical outcome variables, we then used the robust weighted least squares means and variance (WLSMV) adjusted estimator, which does not assume normally distributed variables and provides the best option for modeling non-normal categorical or ordered data (Brown 2015),

to identify the measure’s underlying structure. Once we identified the EFA model, we then tested it in a CFA model.

TABLE IV.5. Quantitative Pre-Testing and Analysis of Survey Measure		
Criterion	Testing Question(s)	Data Analysis
Importance	Does performance on the measure vary by respondent type? How does performance vary by respondent type when different approaches to scoring the measure are applied?	Descriptive analysis (mean, range, outliers) of performance
Factor-analytic structure	How many underlying psychotherapeutic constructs does the measure include? What does the factor structure imply regarding the number of items in the measure?	EFA and CFA
Reliability: Internal Consistency	To what extent do items in each factor measure the same construct?	Alpha statistic
Reliability: Inter-rater	To what extent is there agreement between clinicians, supervisors, and clients in their survey responses?	Agreement using AC1 statistic
Validity	To what extent does the measure distinguish between clinicians who do and do not deliver elements of evidence-based psychotherapy when supervisor ratings are used as the gold standard?	Sensitivity and specificity analyses
Feasibility	On average, how long does it take participants to complete the measure?	Descriptive analysis (means and ranges)
Validity	To what extent does the measure distinguish between clinicians who do and do not deliver elements of evidence-based psychotherapy when supervisor ratings are used as the gold standard?	Sensitivity and specificity analyses
Feasibility	On average, how long does it take participants to complete the measure?	Descriptive analysis (means and ranges)

Confirmatory factory analysis. CFA relies on both empirical and conceptual foundations to guide the specification and evaluation of the factor-analytic model. It is used to test how well a theoretical model fits the data. Unlike in EFA, in CFA the number of factors and the pattern of item-factor loadings are specified in advance. We conducted individual CFAs for the clinician, supervisor, and client samples to further validate the model identified in the EFA. We estimated the models using a Bayes estimator (with the flat empirical priors, 50,000 Monte Carlo Markov Chain chain runs, and two parallel chains), which is less sensitive to sample size (see Heerweg 2014) and does not allow model parameters to fall outside a plausible range (for example,

correlations above one).⁴ We pursued an iterative approach to model-building that included removing the items with low correlation ($r \leq 0.40$) to the latent factor and examining the resulting fit of the model, and made recommendations regarding future revisions to the surveys. We measured the model fit using posterior predictive p-value (PPP; analog of the goodness-of-fit statistics for Bayesian estimator based on the usual chi-square test of the null hypothesis against alternative hypothesis). The general idea behind posterior predictive checking is that there should be little, if any, discrepancy between data generated by the model and the actual data themselves (Kaplan and Depaoli 2012). Hence, p-values greater than 0.05 indicate that the null hypothesis of little discrepancy between the model and the data cannot be rejected and that the model fits the data sufficiently well.

The EFA and CFA models were fitted in *Mplus 7.1* (Muthén and Muthén 1998-2012).

b. Quantitative testing of internal consistency

The internal consistency reliability testing was designed to examine how well the items in each of the five factors correlate to each other and measure the factor's underlying construct. We used the Kuder-Richardson Formula 20 (KR20) and Cronbach's alpha co-efficients. The KR20 is appropriate for categorical items and the Cronbach's alpha for continuous items.

c. Quantitative testing of inter-rater agreement

To assess the extent to which clinicians, supervisors, and clients agreed in their assessment of the clinician's delivery of each survey item, we examined item-level and a weighted average of overall inter-rater agreement using Gwet's Adjusted for Chance-Corrected (AC1) statistic (Gwet 2014). The AC1 is based upon the assumption that the probability of agreement by chance should not exceed 0.50, whereas the probability of chance-agreement for the more traditionally used Cohen's (1960) Kappa can be any value between zero and one.⁵

d. Approaches to establishing performance metrics

For the measure to be useful for quality improvement purposes, stakeholders need metrics to assess performance. There are no clear, established standards for how to score this type of measure. As a first step in developing a measure score, we assessed whether item endorsement varied by beginning, middle, and end of treatment. If there were variation by stage of treatment, our approach to scoring would need to account for it; otherwise, it could overestimate or underestimate a clinician's delivery of evidence-based psychotherapy.

⁴ We also fit the same models using the more conventional WLSMV estimator. The WLSMV relies on large sample theory and assumes a normal distribution. Not unsurprisingly, given the comparatively small sample of clinicians, supervisors, and clients, there were problems identifying the factor model with the WLSMV estimator. Results from both models are presented in Appendix L.

⁵ In preliminary analyses, we calculated inter-rater agreement using the Kappa statistic and observed the "Kappa paradox," where Kappa tends to yield a low value when the raters show high agreement. The AC1 statistic was designed to address the Kappa statistic's limitation. See Appendix M for further information on the Kappa and AC1 statistics.

We conducted Analysis of Variance with post-hoc group comparison to compare the mean scores of each factor identified in the CFA, for each phase of treatment (beginning, middle, and end). No statistically significant differences across phases of treatment were observed. To facilitate comparison across samples (clinicians, supervisors, and clients) and to stabilize variance, factors scores for each domain were standardized to have a mean of zero and a standard deviation of one. Next, we examined the distribution of scores for each domain by respondent type (supervisor, clinician, and (client)). In order to determine potential performance thresholds, we examined various cut-offs (median, mean, inter-quartile range). We selected two thresholds for use in the sensitivity and specificity analyses (described below) -- the median, a lower bound threshold -- and the 75th percentile, an upper bound threshold. Once we created thresholds for each domain, we then created summary scores across all the domains and an overall score. Clinicians who score above these thresholds are classified as delivering evidence-based psychotherapy.

e. Quantitative testing of validity

In addition to gathering feedback from the focus group and site coordinator debriefings on the face validity of the measure (described below), we also attempted to assess the measure's criterion validity by calculating its sensitivity and specificity. For the purposes of these tests, we deemed the supervisor ratings to be the gold standard. In the absence of data from an objective, independent rater, we assumed that supervisors would be the least biased raters and, among supervisors, clinicians, and clients, the raters most trained and experienced in evaluating the performance of clinicians. To calculate specificity and sensitivity, we utilized the performance metrics described earlier and compared supervisor ratings against clinician and client ratings.

6. Approach to gathering stakeholder feedback

In addition to quantitative testing, we gathered feedback on the measure through stakeholder focus groups and site coordinator debriefings. Feedback focused on the importance of the measure to improving quality of care, its face validity, facilitators and barriers to measure testing, the feasibility of implementing the measures (including the burden of data collection), and the usability of the measure results (whether they would be useful for quality improvement efforts). Here, we briefly describe each type of data collection.

Focus groups. In January 2015, we hosted five one-hour telephone focus groups to gather information on the face validity, usability, and feasibility of the measure. Participants represented four types of stakeholders:

- **Clinicians and clinical supervisors.** Focus group participants included eight clinicians and clinical supervisors who had previously completed the survey. Two clinicians who were unable to attend submitted written feedback.
- **Clients.** Participants included four adults in treatment for PTSD who had previously completed the survey. Clients received a \$20 gift card for their participation.

- **Behavioral health organization administrators.** Participants included three administrators from organizations that had participated in pre-testing the survey and one administrator who represented a behavioral health organization that was interested in but unable to participate in pre-testing the survey. One administrator who could not participate submitted written feedback.
- **Health plans and payers.** Participants included eight representatives from four organizations that included two managed behavioral health organizations and two Medicaid managed care organizations.

All questions were designed to answer the main topic areas of usability, feasibility, and validity. We tailored the questions to fit the particular expertise of each type of focus group.

Site coordinator debriefings. To support the site coordinators, Mathematica/NCQA communicated frequently with them. Project staff emailed site coordinators no less frequently than every other day to provide updates on each site's response rates, confirm upcoming therapy session dates, and, if needed, determine if resampling was necessary due to missed therapy appointments or a client terminating therapy. They also held weekly group meetings with the sites to discuss the status of data collection activities and collectively strategize approaches for collecting data.

In addition to the information gathered in the weekly site coordinator meetings, in February and March 2015, we gathered written debriefing information from five sites on ways to improve and streamline data collection processes and on their perceptions of the clinical staff's response to the measure. Site coordinators were also asked to provide their assessment of the measure's face validity, though only some chose to do so. Two sites did not provide any debriefing information.

7. IRB approval and OMB clearance

Prior to the start of data collection, we submitted applications to both the New England Institutional Review Board (NEIRB) and the U.S. Office of Management and Budget (OMB) that outlined the project and its objectives, the proposed study design, sampling and data collection procedures and materials, our security plan, and data analyses. We received approval from the NEIRB on April 29, 2014, and the OMB on May 22, 2014.

8. Processes and procedures to maintain security of data

We implemented the security controls and processes we routinely use on projects that involve sensitive information. Organizations transmitted data to Mathematica via a secure, encrypted Secure File Transfer site that was password-protected. Access to sensitive data was limited to the immediate team and stored on a secure, password-protected network drive. We encrypted data in transit and at rest and will securely destroy any data collected at the end of the project. Hard-copy surveys were mailed or faxed to Mathematica staff for manual entry and stored in a secure, locked file cabinet. We will shred them at the end of the project. These safeguards are consistent with the Privacy Act of 1974, the Computer Security Act of 1987,

Health Insurance Portability and Accountability Act, and the Federal Information Security Management Act of 2002, OMB Circular A-130, and National Institute of Standards and Technology computer security standards.

V. TESTING RESULTS

This chapter summarizes the quantitative and qualitative results of the measure testing. We first present summary statistics on survey administration. We then summarize the factor analyses, internal consistency, inter-rater reliability, measure performance, and sensitivity and specificity. We conclude the chapter with stakeholder feedback.

A. Summary of Survey Administration

Survey mode. Eighty-nine percent of clinicians, 63 percent of supervisors, and 37 percent of clients completed the survey via the web (Table V.1). The mode of survey completion varied by site. For example, in Site B, clients were provided the option of completing the survey immediately following the therapy session using the site's iPads. All of the clients at this site completed the survey using the web; 63 percent of the clients opted to complete it before leaving the site following their session (data not shown). Conversely, 100 percent of clients at Site D elected to complete the survey on paper.

Length of time to complete the survey. On average, clinicians completed the web survey in 8 minutes and supervisors and clients in 10 minutes (Table V.1).⁶ We excluded from these calculations 17 cases where the response times were greater than one hour. It is likely that these outlying values reflect individuals who started the survey, saved their responses, and completed the survey at a later time.

Length of time between therapy session and survey completion. To reduce recall bias, clients and clinicians were asked to complete the survey within 24 hours of the therapy session, and supervisors were asked to complete it within 24 hours of their review of the session. Table V.1 suggests that, on average, clinicians and clients did not complete the survey within this 24-hour window. The average number of days between when the therapy session occurred and when clinicians and clients completed the survey was 9.6 days (range 0-127 days) and 2.0 days (range 0-12 days), respectively. We do not have information on when the supervisors began their review of the therapy session; however, the average length of time between the occurrences of the therapy session and when supervisors completed the survey was 20 days (range 0-102 days).

Multiple factors may contribute to the length of time between the occurrence of the therapy session and survey completion. Conversations with site coordinators indicate that in some cases the length of time may be an artifact of clinicians and supervisors saving their survey responses but not actually clicking the "submit" button to transmit them. If the survey were to undergo future testing, revisions to the web version could provide additional prompts to submit the survey upon completion. Additionally, some site coordinators indicated that supervisors conducted weekly supervision and reviewed session tapes in batches; this may contribute in part to the delayed completion of the surveys. It is also likely that the data may accurately reflect the time

⁶ We are unable to calculate length of time to complete paper surveys.

needed for clinicians and supervisors to complete the survey, in which case, further investigation is needed into recall bias and the accuracy of the data when the survey is completed days and sometimes weeks after the therapy session occurred. Further investigation may also be needed into the organizations' capacity to complete this type of quality measure, and into the resources -- and perhaps changes in internal processes -- needed to facilitate more timely survey completion. In considering processes that facilitate data collection, regular reminders to staff to complete the survey appear key. The coordinators at Sites C and E were especially responsive to Mathematica alerts to remind staff of outstanding surveys, and these sites have comparatively shorter survey completion times. Routine reminders to clinicians and supervisors to complete the measure may be an important part of collecting the data in a timely way.

TABLE V.1. Summary of Survey Administration: Modes and Completion Times						
		Number of Completed Surveys	Percentage Web-Based Complete (n)	Percentage Paper-Based Complete (n)	Average Number of Minutes to Complete the Survey (range)***	Average Number of Days from Therapy Session Start Date to Survey Completion (range)****
Total	Clinicians	96	89% (85)	11% (11)	8 (2-56)	9.9 (0-127)
	Supervisors	97	63% (76)	37% (21)	10 (2-52)	19.6 (0-102)
	Clients	78	37% (29)	63% (49)	10 (3-30)	1.9 (0-12)
Site A	Clinicians	34	100% (34)	0% (0)	9 (3-56)	14 (0-127)
	Supervisors	34	100% (34)	0% (0)	10 (2-47)	27 (0-76)
	Clients	22	23% (5)	77% (17)	6 (3-11)	3.8 (0-7)
Site B	Clinicians	10	100% (10)	0% (0)	9 (2-13)	20 (0-72)
	Supervisors	10	100% (10)	0% (0)	7 (3-20)	24 (0-102)
	Clients	8	100% (8)	0% (0)	7 (4-11)	2.5 (0-9)
Site C	Clinicians	15	100% (15)	0% (0)	6 (2-13)	0.5 (0-4)
	Supervisors	15	100% (15)	0% (0)	9 (4-52)	8 (2-20)
	Clients	14	7% (1)	93% (13)	15 (15)	0 (0)
Site D	Clinicians	12	58% (7)	42% (5)	10 (2-23)	6.6 (0-14)
	Supervisors	14	100% (15)	0% (0)	12 (4-23)	13.6 (0-51)
	Clients	13	0% (0)	100% (13)	NA	NA
Site E	Clinicians	18	89% (16)	11% (2)	6 (3-14)	4.4 (0-29)
	Supervisors	17	18% (3)	82% (14)	9 (9)	2.7 (0-7)
	Clients	17	82% (14)	18% (3)	11 (6-30)	1 (0-12)
Site F	Clinicians	7	43% (3)	57% (4)	8 (6-9)	8.3 (1-22)
	Supervisors	7	0% (0)	100% (7)	NA	NA
	Clients	4	25% (1)	75% (3)	26 (26)	1 (1)

* Paper-based completes are excluded, because the information is not available.
 ** Durations over one hour were excluded (17 cases out of 191 total), as it is likely that these participants completed the survey in more than one sitting.
 *** Days calculated are calendar days.

Item-Level Missing Information. Most participants entered a response for each survey item. On the clinician survey, eight items had missing information and the missingness ranged from 0-2 percent (Appendix G, Table G.1). On the supervisor survey, 28 items had missing information; the level of missing information ranged from 0-3 percent (Appendix G, Table G.2). On the client survey, 30 items had missing information, which ranged from 0-6 percent (Appendix G, Table G.3).

B. Exploratory Factor Analysis

To identify the underlying factor structure of the survey, we fit a series of EFA models with varying numbers of latent factors (5, 6, 7, and 8). We examined the models' statistical fit

and how well they corresponded to our theoretical understanding of the underlying constructs of evidence-based psychotherapy for PTSD.

According to the model fit statistics (Appendix H, Table H.1), all four of the EFA models represented the underlying data structure very well, suggesting that from a statistical standpoint any of these models could inform the CFA. We then examined the factor structures for parsimony and clinical meaning. The five-factor model provided the most parsimonious solution with the least number of significant cross-loadings. This solution was also the most interpretable based on constructs identified during the measure development stage. For these reasons, we retained the five-factor model (see Table V.2) for further validation at the CFA stage.

In grouping items into factors, we considered items with factor scores of 0.40 or above. If an item cross-loaded on multiple factors, we assigned it to the factor where it had the highest loading and/or for which other factors related to the item also scored highly. Below, we describe the factor groupings and the labels we assigned to each factor.

- **Factor 1: Structuring and conducting the therapy session.** Ten items compose Factor 1 and include aspects of treatment such as creating an agenda, setting treatment goals with the client, soliciting client feedback on treatment, and being directive.
- **Factor 2: Psychoeducation and therapeutic techniques.** Fifteen items make up Factor 2. The majority of items are therapeutic techniques (that is, cognitive restructuring, Socratic method, imagining the traumatic event) and psychoeducation (providing education about symptoms and/or the traumatic event).
- **Factor 3: Therapeutic alliance.** Three items from the therapeutic alliance measure make up Factor 3.
- **Factor 4: Assessment.** Two items on assessment loaded on this factor.
- **Factor 5: Homework.** All six of the items that loaded on this factor are related to assigning, reviewing, and encouraging homework completion.

Five items were not statistically significantly associated with any of the factors. These items included the suicide risk assessment questions, use of Socratic questioning, the facilitation of alternate hypotheses, and one question on the clinician's time management.

TABLE V.2. The Five-Factor EFA Solution					
Clinician Survey Item Number	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
12.a. IMAGINE		0.699*			
12.b. WRITE		0.783*			
12.c. OTHER SOCRAT	-0.266*	0.867*			
12.d. REAL		0.578*			
26. TRUST			0.750*		
24. CONFIDENT			0.774*		
25. LIKES			0.818*		
1. AGENDA	0.790*				
2. REVIEW AGENDA	0.719*	0.376*		-0.290*	
3. BACKGROUND	0.307*	0.676*			
4. EXPECTATIONS	0.696*				
5. GOALS	0.647*				
10. IDENTIFY		0.670*			
7. COG RESTRUC		0.469*			
8. SOCRAT					
9. FACILITATE					
10. OTHER IDENTIFY		0.556*			
11. TECHNIQUES		0.633*			
13. DISCUSS	0.226*	0.549*			
14. STRUGGLE					
15. DIRECTIVE	0.853*				
16. TX FEEDBACK	0.604*	0.233*			
17. TH FEEDBACK	0.431*				
18. ASSIGN					0.769*
19. REVIEW INSTRUC					0.722*
20. ADDRESS					0.949*
21. SOLUTION					0.749*
22. REVIEW HMWK					0.925*
23. ENCOURAGE					0.874*
27.a. EVER SUIC					
27.b. TODAY SUIC	0.269*				
28.a. EVER USE SUIC					
28.b. TODAY USE SUIC					
29.a. EVER INSTRU				0.614*	
29.b. TODAY INSTRU	0.466*			0.750*	
30.a. EVER SYMP EDU		0.699*		-0.393*	
30.b. TODAY SYMP EDU		0.836*			
31.a. EVER TRAUMA ED		0.574*			
31.b. TODAY TRAUMA ED		0.768*			-0.254*
32.a. EVER OUTLINE	0.516*				
32.b. TODAY OUTLINE	0.685*	0.369*			
12. OVERALL TECHNIQUES		0.737*			
* Factor loadings not significant at $p \leq 0.05$ were excluded from the table to facilitate interpretation of the results.					

C. Confirmatory Factor Analysis

To further refine the scales identified in the EFA, we conducted CFAs on the five-factor model separately for the clinician, supervisor, and client samples. The CFA models fit the data well and had a similar factor structure across the different respondents (Appendix H, Table H.2), suggesting that the instrument may function similarly across the three types of respondents. Detailed CFA results by respondent type are available in Appendix I. A summary of the commonalities and differences in the factor structures across the samples is below:

- **Factor 1: Structuring and conducting the therapy session.** The number of items that compose Factor 1 varies by respondent type. Across the three samples, five items related to agenda setting, goals, treatment process and expectations, and treatment feedback make up this factor. In the clinician and supervisor surveys, this factor also comprises reviewing agendas and being directive. Outlining the treatment process and symptom assessment also loaded on Factor 1 in the clinician and client surveys.
- **Factor 2: Psychoeducation and therapeutic techniques.** The items that compose Factor 2 are nearly identical across the three samples and, as previously described, focus on therapeutic techniques such as the use of Socratic questioning and cognitive restructuring.
- **Factor 3: Therapeutic alliance.** The three therapeutic items compose Factor 3 across all three samples.
- **Factor 4: Assessment.** This factor has only one item, suicide risk assessment “today,” shared between the three samples. Each paired sample (clinician/client, clinician/supervisor, client/supervisor) has common items that make up this factor. The items include therapeutic techniques and additional assessment questions.
- **Factor 5: Homework.** The items that compose Factor 5 are nearly identical. It has four common items across the three samples and five common items between the clinicians and supervisors.

Summary. Taken together, the EFA and CFA results suggest that the survey items measure constructs relate to the delivery of psychotherapy for PTSD. For further instrument development, we recommend analyzing whether core items that are consistent across all three samples are sufficient to capture the corresponding latent factors without sacrificing the reliability of a scale. This could help to shorten the measurement instrument and decrease the burden for respondents while retaining essential measurement properties. We also recommend considering modifications to the fourth factor, which only has one item shared by all three samples and which also has the lowest scale reliability of all five factors.

D. Internal Consistency Results

According to our KR20 analysis, the internal consistency of four out of five latent constructs is between 0.70 and 0.90 (Table V.3; details shown in Appendix J), which is in the “good” to “very good” range (Nunnally and Bernstein 1978). The internal consistency of Factor 4, suicide assessment, is between 0.54 and 0.69, which suggests some items may need revision. On average, we observed the highest reliability across all domains in the supervisor sample, followed by the clinician and client samples.

TABLE V.3. Internal Consistency Results by Factor and Respondent					
Respondent	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Clinician	0.78	0.83	0.82	0.58	0.81
Supervisor	0.88	0.89	0.85	0.69	0.81
Client	0.77	0.77	0.82	0.54	0.90

E. Inter-Rater Agreement Results

Inter-rater reliability assesses the extent to which clinicians, supervisors, and clients agreed on whether the clinician delivered the survey element. We used the AC1 statistic, a measure of agreement adjusted for chance, to quantify agreement for the overall survey and at the item level.⁷

Inter-rater agreement between clinicians, supervisors, and clients. All three raters completed the survey on 76 therapy sessions and at least two raters completed it on 97 therapy sessions. The weighted agreement for the whole survey ranged from 0.39 to 0.58 across different rater pairs (Table V.4), which is considered fair-to-moderate agreement (Gwet 2014). Supervisors and clinicians had the highest weighted inter-rater agreement; supervisors and clients and clinicians and clients had comparable inter-rater agreement.

TABLE V.4. Inter-Rater Reliability for Clinicians, Supervisors, and Clients				
Raters	AC1	SE	CI	Significance Level
Supervisor-clinician-client	0.43	0.005	(0.34-0.53)	< 0.01
Supervisor-clinician	0.58	0.04	(0.51-0.65)	< 0.01
Supervisor-client	0.39	0.07	(0.25-0.54)	< 0.01
Client-clinician	0.39	0.07	(0.26-0.51)	< 0.01
NOTE: AC1 values above 0.80 suggest high agreement; 0.61-0.80 substantial agreement, 0.41-0.60 moderate agreement, 0.21-0.40 fair agreement, and 0-0.20 slight agreement.				

In addition to calculating inter-rater agreement for the whole measure, we also calculated it at the item level. Across the three raters, item percentage agreement ranged from 39 percent to 90 percent and the AC1 statistic ranged from -0.09 to 0.86 (Table V.5). Items for which there was only slight agreement included two homework-related items, one therapeutic technique item, and one item on managing the therapy session. Similar trends occurred when examining item-level agreement between each rater pair (clinicians/supervisors, clinicians/clients, supervisors/clients) with high agreement in ratings of some survey items and low agreement in others (see Appendix K).

⁷ At the beginning of data collection, the supervisors in one site mistakenly completed 22 surveys based on review of the clinician's case notes instead of audio tape review. We calculated inter-rater reliability with and without the 22 surveys. Most agreement measures were negligibly affected by the exclusion, although one item, with regard to the therapist struggling to manage time, did dramatically decrease, from 0.81 to -0.67. Overall, these results indicate that completion of the supervisor survey did not create significant bias in the results of our inter-rater agreement analysis.

TABLE V.5. Item-Level Inter-Rater Agreement Between Clinicians, Supervisors, and Clients				
	Overserved Agreement	AC1	CI	Significance Level
Did you and your therapist discuss an agenda or plan for your session?	85.48%	0.42	< 0.01	(0.24-0.60)
Did your therapist talk about or check-in on your expectations of how therapy will go?	62.50%	0.38	< 0.01	(0.20-0.56)
Did your therapist work with you to set goals you both agreed on?	65.45%	0.52	< 0.01	(0.35-0.69)
Did your therapist help you become aware of or realize feelings, views or thoughts in your life that have been influenced by your traumatic experience? These might include feelings, views, or thoughts about being safe in the world, the presence of danger, trust, and self-esteem.	76.71%	0.69	< 0.01	(0.57-0.82)
Did your therapist ask you several direct questions to make you think critically about or examine your thoughts, feelings, or beliefs? For example, your therapist might ask: <ul style="list-style-type: none"> • How do you know this? Can you give me an example? • What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? • How did you come to this conclusion? What evidence do you have to justify this? 	58.62%	0.48	< 0.01	(0.31-0.64)
Did your therapist offer other ways of thinking about your issues (e.g., problem areas or areas you want to work on) related to the trauma? For example: <ul style="list-style-type: none"> • Thought: "I can't trust anyone." • Thought suggested by therapist: "Some people can't be trusted, but there are other people who are trustworthy." 	64.81%	0.41	< 0.01	(0.23-0.60)
Did you and your therapist discuss people, events, or places you now avoid or stay away from because of your traumatic experience? For example, someone in a car accident might avoid driving on the freeway.	61.67%	0.20	< 0.01	(0.01-0.39)
Did your therapist do any of the following things to help you deal with fear, anxiety or things you now avoid because of your trauma? <ol style="list-style-type: none"> a) Ask you to imagine or retell your traumatic experience for longer than 10 minutes. b) Ask you to write about your traumatic experience. c) Ask you questions to make you think critically about or examine your thoughts, feelings, or beliefs related to your fear, anxiety, and avoidance of things (i.e., "How do you know this? Can you give me an example?"). d) Ask you to do real world experiments like visiting a place related to the traumatic experience for longer than 10 minutes. 	49.30%	0.22	< 0.01	(0.06-0.38)
After you described your traumatic experience, did you and your therapist discuss the details of what happened to you, how it impacted your life, or your emotions about the event?	62.86%	0.19	< 0.01	(0.03-0.35)
Did your therapist make good use of your session time today?	76.19%	-0.09	< 0.01	(-0.23-0.05)
Did your therapist ask for your opinion on how your treatment is going?	66.07%	0.50	< 0.01	(0.34-0.66)
Did your therapist ask for feedback on how she/he is doing in helping you recover from your PTSD?	45.10%	0.25	< 0.01	(0.07-0.44)

TABLE V.5 (continued)				
	Overserved Agreement	AC1	CI	Significance Level
Did your therapist assign homework or practice assignments (to be completed by the next session) to work on your PTSD symptoms or problem areas?	58.82%	0.33	< 0.01	(0.16-0.50)
Did your therapist make sure you understood how to complete your homework for the next session?	65.52%	0.45	< 0.01	(0.29-0.61)
If you had problems completing your previously assigned homework, did your therapist work with you to come up with solutions to these problems?	66.07%	0.09	< 0.01	(-0.08-0.27)
Did your therapist review and discuss your homework from the previous session?	54.55%	0.17	< 0.01	(-0.01-0.35)
When reviewing the homework from the previous session, did your therapist encourage or provide you with constructive feedback?	60.00%	0.22	< 0.01	(0.03-0.40)
My therapist and I have built mutual trust.	10.67%	0.85	< 0.01	(0.80-0.90)
I am confident in my therapist's ability to help me.	9.33%	0.76	< 0.01	(0.70-0.83)
I believe my therapist likes me as a person.	17.33%	0.86	< 0.01	(0.81-0.91)
Has your therapist ever asked you if have had thoughts about committing suicide?	90.00%	0.79	< 0.01	(0.68-0.89)
During this session, did your therapist ask you if you had thoughts about committing suicide?	39.39%	0.61	< 0.01	(0.47-0.76)
Has your therapist ever asked you to answer questions about your PTSD symptoms? This might include completing a form before or after therapy.	78.69%	0.28	< 0.01	(0.12-0.43)
During this session, did your therapist ask you about your PTSD symptoms? This might include completing a form or survey before or after therapy.	61.54%	0.44	< 0.01	(0.29-0.59)
Has your therapist ever provided information about PTSD and PTSD symptoms?	84.75%	0.86	< 0.01	(0.77-0.95)
During this session, did your therapist provide information about PTSD and PTSD symptoms?	60.94%	0.26	< 0.01	(0.09-0.43)
Has your therapist ever provided with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your viewpoints and beliefs.	68.00%	0.48	< 0.01	(0.31-0.65)
During this session, did your therapist ever provide you with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your view points and beliefs.	56.67%	0.22	< 0.01	(0.06-0.38)
Has your therapist ever explained how your particular treatment will work?	78.18%	0.77	< 0.01	(0.65-0.89)
During this session, did your therapist explain how your particular treatment will work?	67.21%	0.21	< 0.01	(0.05-0.36)

Implications for survey revisions. Although there was high agreement between raters for several survey items, the inter-rater agreement results suggest that several items may benefit from further investigation and potential revision. Examples of items with low agreement and/or poor AC1 values include:

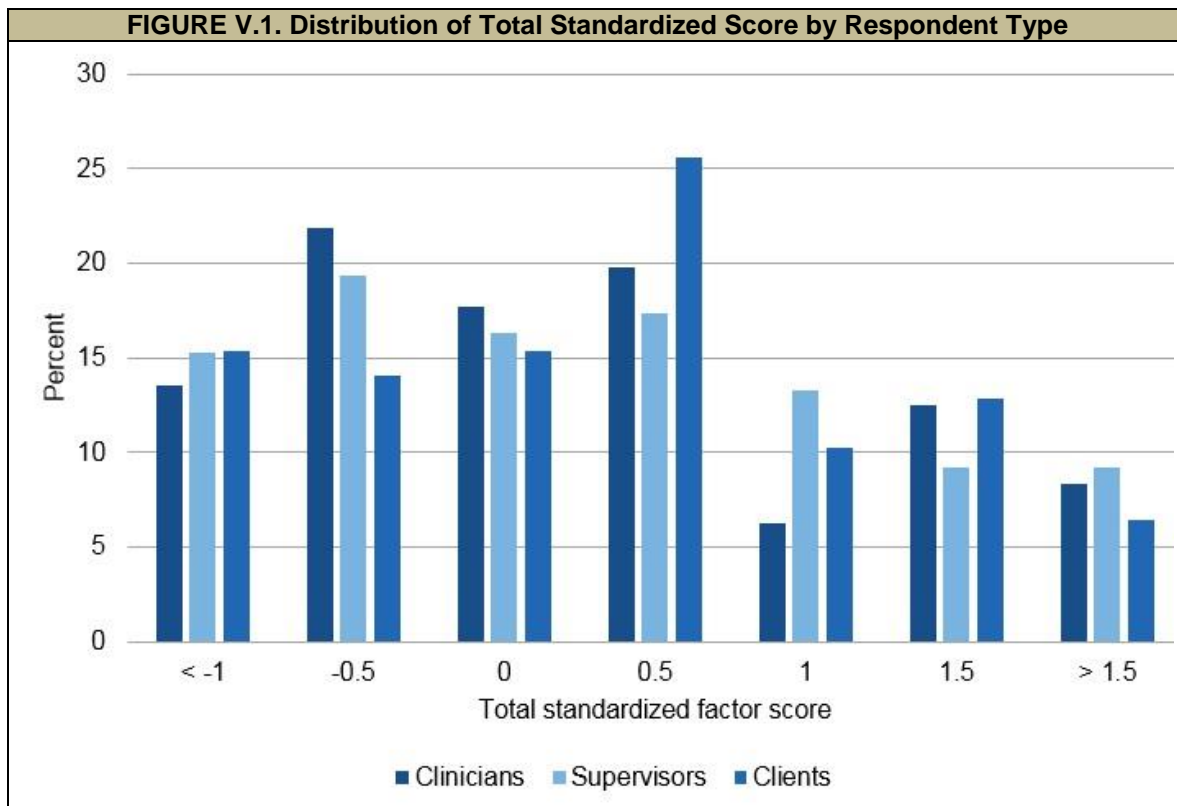
- Two questions regarding Socratic discussion methods.
- Two questions about therapeutic techniques to deal with avoidance.

- One question about emotional reprocessing regarding the emotions surrounding the traumatic event.
- One question regarding the psychoeducation about the nature of the traumatic event.

It is possible that these and other items with low agreement could be revised by further simplifying the questions or providing more detailed examples; however, further cognitive interviewing may be needed to better understand how stakeholders interpret them. Alternatively, the items may need to be deleted.

F. Approach to Creating a Measure Score

In order for a measure to be useful for performance and accountability purposes, the measure must discriminate performance and there must be a mechanism for scoring it to identify individuals who delivery evidence-based psychotherapy. As an initial approach to developing a measure score, we created standardized factor scores for each of the five factors identified in the factor analyses. The scores were standardized to have a mean of zero and a standard deviation of one. A total standardized score was also created using the same method. As depicted in Figure V.1, the distribution in total scores varies for each of the three respondent types.



Next we examined approaches to establishing measure thresholds that could be used to identify clinicians who deliver evidence-based psychotherapy. We examined four thresholds: the

mean, median, mean plus one standard deviation, and the 75th percentile. We selected two thresholds -- the median and 75th percentile -- as more conservative and liberal estimates of measure performance for further investigation. In the subsequent section, we describe the measure's performance when using these thresholds.

G. Results of Sensitivity and Specificity Analyses

To begin to understand the measure's validity, we calculated its sensitivity and specificity. For the purposes of this investigation, sensitivity is defined as the proportion of clinicians identified by clients or the clinicians themselves as high performers in the delivery of evidence-based psychotherapy when compared to supervisor scores. Specificity, in contrast, is the proportion of clinicians identified as low performers in the delivery of evidence-based psychotherapy. We compared clinician and client scores to the supervisor scores, which for the purposes of these analyses, we treated as the gold standard. We examined the implications for the measure's sensitivity and specificity using two thresholds, the median (P5) and above the 75th percentile (P75) to determine high and low delivery of evidence-based psychotherapy.

Table V.6 summarizes the sensitivity and specificity results. For supervisors and clinicians, the sensitivity rate ranged from 0.32 to 0.78 across the factors. The specificity rate ranged from 0.51 to 0.88. For supervisors and clients, the sensitivity rate was 0.22-0.61 and the specificity rate was 0.49-0.81 (Table V.6).

Based on these preliminary findings, the P50 (median) threshold appears to better discriminate performance than the more stringent P75 threshold. This threshold obtained consistently higher values for sensitivity and specificity in supervisor-clinician pairings when compared to the P75 threshold.

In both supervisor-clinician and supervisor-client pairings, the P75 threshold demonstrated higher specificity. However, in supervisor-client pairings, the sensitivity values with the P75 were quite low compared to those observed among the clinicians at the same threshold, suggesting a differential performance with the instrument between respondents. The observed differences in performance across pairings suggest a need to further evaluate the instrument to identify the optimal threshold for each respondent type.

When thinking about measure implementation, it is important to note there may be instances where a supervisor is not the gold standard. For example, supervisors may treat too few patients to serve as experts in the delivery of evidence-based psychotherapy or they may not be trained in cognitive behavioral approaches--which the measure largely draws upon--and therefore, may not be best positioned to identify a clinician's use of these techniques. In Chapter VI, we discuss next steps for further assessing the measure's validity.

TABLE V.6. Results of Sensitivity and Specificity Analyses								
	Comparison of Supervisor and Clinician Scores				Comparison of Supervisor and Client Scores			
	Specificity P50 Threshold	Sensitivity P50 Threshold	Specificity P75 Threshold	Sensitivity P75 Threshold	Specificity P50 Threshold	Sensitivity P50 Threshold	Specificity P75 Threshold	Sensitivity P75 Threshold
Factor 1: Structuring and conducting the Session	0.78	0.78	0.81	0.45	0.49	0.50	0.81	0.32
Factor 2: Psychoeducation and therapeutic techniques	0.51	0.50	0.78	0.32	0.56	0.56	0.76	0.26
Factor 3: Therapeutic Alliance	0.63	0.64	0.82	0.48	0.50	0.51	0.75	0.22
Factor 4: Suicide assessment	0.61	0.63	0.85	0.57	0.56	0.58	0.79	0.37
Factor 5: Homework	0.63	0.63	0.80	0.42	0.62	0.61	0.76	0.32
Overall score	0.73	0.74	0.85	0.50	0.61	0.62	0.76	0.32

H. Stakeholder Feedback

In January 2015, we held four discussion groups with clinicians and supervisors, clients, site administrators, and health plans and payers to gather feedback on the measure’s importance, face validity, usefulness, and feasibility. During this time, we also gathered feedback from site coordinators. Below, we summarize key themes identified across the discussions. Given overlapping themes in the feedback provided, we include information learned from the site coordinator briefings in this section.

Importance. Stakeholders agreed on the importance of improving the quality of PTSD care. Perceptions regarding this measure’s importance varied. Health plans indicated a strong preference for outcomes measures and indicated that additional process measures have little utility in improving quality of care.

Validity. Perceptions regarding the measure’s face validity were mixed.

- **Measuring true quality of treatment.** Several clinicians, administrators, and health plan/payer representatives, and site coordinators suggested the measure was too narrowly focused on cognitive behavioral approaches and did not cover the range of (perceived appropriate) treatments for adults with PTSD. Others felt the survey items reflected the true quality of evidence-based treatment.

Usability. Stakeholders had mixed opinions regarding the usefulness of the measure.

- **Usefulness.** Stakeholders agreed the measure would be useful for training and continuing education purposes; however, there was a lack of consensus regarding its usefulness for quality improvement. Clients, administrators, and some clinicians suggested the measure would also be useful for accountability and quality improvement; however, health plan/payer representatives uniformly agreed the measure would not be of

use. Given the relatively small proportion of their beneficiaries who are in treatment for PTSD and the emphasis on the development of outcome measures, the health plan and payer group representatives would not find the measure useful.

- **Service setting.** Stakeholders suggested the measure could be useful for outpatient clinics, the VA, day hospital programs, and PTSD Centers of Excellence. Site administrators and health plan/payer representatives did not perceive the measure as being useful for health plans.
- **Unintended consequences.** Some clients suggested that survey completion might unintentionally result in a potential confrontation between clients and clinicians. Participants in the client group offered a scenario in which a client indicated that his clinician did not provide most of the items on the survey. In this scenario, participants worried that the client's survey responses would be shared with the clinician and influence the nature of the subsequent session. In order to avoid this potential scenario, some clients suggested making the survey anonymous to the clinician. In contrast, others from the client group said that they would take the opportunity to influence their course of treatment. This group of clients stated that if their clinician were not receptive to the feedback, they would discontinue treatment and find a new clinician. Some clients with good relationships with their clinicians indicated this was an unlikely scenario.
- **Other concerns.** Administrators stated that recording or directly observing therapy sessions could hinder clients' willingness to complete the survey or make them wary of speaking freely during a session out of fear of repercussions. Some clients also expressed concerns about unintended consequences and specifically about how clients might react if, based on the survey, they felt the clinician was not delivering quality care.

Feasibility. With the exception of clients, all stakeholder groups expressed concerns regarding the measure's feasibility.

- **Prioritization of surveys.** All stakeholder groups suggested it would be too resource intensive to utilize all three versions of the survey. Given the time, resources, and (in some cases) changes in supervision processes that would be required, none of the groups selected the supervisor version of the survey for administration. Health plan/payer representatives indicated a preference for the client version. Site administrators and clinicians indicated they would choose either the clinician or the client version.
- **Survey length.** Health plan/payer representatives and some site coordinators felt the survey was too long.
- **Survey mode.** Stakeholder feedback on the feasibility of implementing web-based surveys varied. Some stakeholders found it convenient and time-saving; others experienced challenges in navigating the online survey and indicated that many clients do not have reliable Internet access.

- **Coordination.** Administrators and site coordinators expressed concerns regarding the feasibility of coordinating the data collection effort, particularly in drawing the sample and providing reminders to the participants to complete the survey. The administrator from one site also indicated concerns regarding the resources required to translate the materials into other languages.

I. Summary of Site Coordinator Debriefings

Site coordinators provided written feedback at the end of data collection. The following are topic areas of the types of feedback we received:

Technological challenges. Some respondents had difficulty using the online survey links, whereas others found the links to be user-friendly. Both staff and consumer respondents at some sites found it easier to complete paper copies of the surveys.

Survey questions. Some sites found the questions to be too targeted to CPT and prolonged exposure therapy, which could skew the results, since not all clients sampled received that type of treatment. Some site coordinators heard from supervisors and clinicians that the survey questions were a useful reminder to stick to evidence-based treatment and to utilize certain tactics in all sessions. Some clinicians found the questions to be too generic.

Participant hesitance:

- **Clients.** Many sites struggled with client hesitance about participating in the study. Some were initially wary of having an observer present during their session or having their session audio taped, but coordinators said that most clients forgot about the observer or audio tape by the end of the session. However, sites reported that, overall, many clients were excited to participate and, despite initial hesitance, were willing to participate if it could help others receive high quality care in the future.
- **Clinicians.** Many sites reported that most of the clinicians were cooperative and excited to participate. However, some were hesitant about being observed and some were unclear about what would happen with the results of the survey.

Scheduling and time commitment:

- Many sites reported that tracking client appointments and client absences and re-schedules was challenging.
- One site did not fully understand the supervisor time commitment (to observe or review every selected session in its entirety) when they originally agreed to participate.

- Some sites did not fully understand the site coordinator time commitment, and found that the role was too much for one person. One site mentioned that internal logistics were challenging.
- One site would have liked more time for data collection.

VI. CONCLUSIONS AND NEXT STEPS

The development of a measure of the delivery of evidence-based psychotherapy has the potential to improve the quality of care for adults with PTSD. We made promising strides in creating the foundation of such a measure; however, a significant amount of additional work is needed to develop a final measure that can be used for accountability purposes. Below, we provide overarching conclusions and recommended next steps.

- **Additional input.** Although there was support for use of the measure in training and education, support for using it for accountability purposes was limited. Additional input from a larger group of stakeholders regarding the measure's use for internal quality improvement and the circumstances under which it would be useful would inform the next stages of measure development.
- **Further revisions.** Our analyses suggest that the survey assesses important underlying constructs associated with the delivery of evidence-based treatment for PTSD and that many survey items produce significant agreement across the three raters. The analyses also suggest that several items need refinement. For example, items with low inter-rater agreement and/or low internal consistencies may be candidates for deletion. Items with significant cross-loadings and moderate agreement could need revision. The surveys should be revised further, with additional cognitive testing and stakeholder input conducted on the refinements.
- **Further investigation of feasibility.** Several stakeholders expressed concern regarding the measure's feasibility. Refinement to the survey items may result in a shorter measure that takes less time to complete, which should improve the feasibility of using it. In addition, it would be useful to have additional information from a larger group of stakeholders regarding topics such as preferred survey mode (including mobile technology applications), the available infrastructure to support the measure, and approaches to automating aspects of site coordination.
- **Further development of the measure for broader application.** The factor analyses results identified therapeutic constructs that are likely relevant in the delivery of psychotherapy for conditions other than PTSD. The measure could be refined and further tested to create modules that broadly apply to the delivery of psychotherapy.
- **Examine inter-rater reliability and factor structure with revised items and larger sample.** Once the survey items have been refined, additional work will be needed to test whether the refinements improve inter-rater agreement and the factor structure. The goal of our current project was to pre-test this instrument. A pilot test with a larger sample offering increased diversity in sites, clinicians, and clients would increase the external validity of the measure.

- **Examine other scoring methods.** Our current thresholds for high and low delivery of evidence-based psychotherapy yielded positive results in terms of specificity and sensitivity. After item refinement, these scoring methods should be verified and compared to other possible methods of scoring. For example, contextual scoring may be beneficial, as it would allow clinicians flexibility in deviating from a treatment plan for appropriate reasons (for example, in cases where a clinician did not use an expected set of therapeutic elements, because he or she had to help a client manage suicidal ideation).
- **Additional validity testing.** Additional psychometrics are needed to validate this measure. The use of an external, independent rater (not associated with the site) to serve as the preferred gold standard is important. To assess the measure's predictive validity, information on patient outcomes (for example, symptom improvement, quality of life, and functioning) is critical.

The measure developed under this project has the potential to address significant gaps in quality of PTSD care. Additional work is needed to further prepare it for implementation on a larger-scale basis and to better understand the groups and situations where the measure will be most useful.

REFERENCES

- Agency for Healthcare Research and Quality (AHRQ). "Background Report for the Request for Public Comment on Initial, Recommended Core Set of Children's Healthcare Quality Measures for Voluntary Use by Medicaid and CHIP Programs." Available at <http://www.ahrq.gov/chipra/corebackground/corebackapa8.htm>. Accessed February 7, 2011.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: Author.
- Brook, R.H., R.E. Park, M.R. Chassin, D.H. Solomon, J. Keeseey, and J. Kosecoff. (1990). Predicting the Appropriate Use of Carotid Endarterectomy, Upper Gastrointestinal Endoscopy, and Coronary Angiography. *New England Journal of Medicine*, 17, 1173-77.
- Brown, T.A. (2015). *Confirmatory Factor Analysis for Applied Research*. New York, London: Guilford Press.
- Chorpita, B.F., and E.L. Daleiden. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology*, 77, 566-579.
- Chorpita, B. F., E. Daleiden, and J.R. Weisz. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research*, 7, 5-20.
- Fitch, K., S.J. Bernstein, M.D. Aguilar, B. Burnand, J.R. Lacalle, and P. Lazaro. *The RAND/UCLA Appropriateness Method User's Manual*. RAND Corp, 2001.
- Gwet, K. (2014). *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters*, 4th ed. Advanced Analytics, LLC.
- Heerwegh, D. (2014). Small Sample Bayesian Factor Analysis. Paper presented at PhUSE annual conference (London 12-15 October 2014).
- How Common is PTSD? (2007, July 5). U.S. Department of Veteran Affairs National Center for PTSD. Retrieved September 3, 2013, from <http://www.ptsd.va.gov/public/pages/how-common-is-ptsd.asp>.
- Institute of Medicine (IOM). *Treatment of Posttraumatic Stress Disorder: An Assessment of the Evidence*. Washington, DC: The National Academies Press, 2008.
- Institute of Medicine. *Treatment for Posttraumatic Stress Disorder in Military and Veteran Populations*. Washington, DC: National Academies Press, 2012.

- Kaplan, D. and S. Depaoli. (2012). Bayesian structural equation modeling. In R. Hoyle (ed.), *Handbook of Structural Equation Modeling*. (pp 650-673), New York, NY: Guilford Publications, Inc.
- Kessler, R.C., P.A. Berglund, O. Demler, R. Jin, K.R. Merikangas, and E.E. Walters. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry*, 62(6), 593-602. PubMed Abstract, Lifetime prevalence of DSM-IV disorders by sex and cohort Table 1.
- Kessler, R.C., W.T. Chiu, O. Demler, K.R. Merikangas, and E.E. Walters. (2005). Prevalence, severity, and comorbidity of twelve-month DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry*, 62(6), 617-627. PubMed Abstract, Erratum, 12 month prevalence of DSM-IV disorders by sex and cohort Table 2.
- Mellman, T.A., R.E. Clark, and W.J. Peacock. (2003). Prescribing patterns for patients with posttraumatic stress disorder. *Psychiatric Services*, 54, 1618-21.
- Muthén, L.K. and B.O. Muthén. (1998-2012). *Mplus User's Guide*, 7th ed. Los Angeles, CA: Muthén & Muthén.
- Nunnally, J.C., and I.H. Bernstein. (1994). *Psychometric Theory*, 3rd ed. New York, NY: McGraw-Hill.
- Seng, et al. (2009). Prevalence, trauma history, and risk for posttraumatic stress disorder among nulliparous women in maternity care. *Obstetrics and Gynecology*, 114, 839-947.
- Tuerk, P.W., B. Wangelin, S.A. Rauch, C.E. Dismuke, M. Yoder, H. Myrick, A. Eftekhari, and R. Acierno. (2012). Health service utilization before and after evidence-based treatment for PTSD. *Psychological Services*.

APPENDIX A. PTSD TAG1 PRESENTATION

Development of Quality Measures for Post Traumatic Stress Disorder Technical Advisory Group

March 1, 2012

MATHEMATICA
Policy Research

TAG AGENDA

- Introductions
- Conflict of Interest
- Background Information
- Measure Development Process
- Measure Concepts
- Wrap-Up
- Next Steps

MATHEMATICA
Policy Research

Project Overview

- ASPE and NIMH are funding this project
- Purpose is to develop quality measures of care provided to adults with PTSD treated in ambulatory settings
- The focus is on care provided to adults *diagnosed* with PTSD treated *outside* of the VA
- Mathematica is the prime contractor and has a subcontract with NCQA

Project Goals

- Goals in the next year
 - Develop a foundation for measuring quality of PTSD treatment
- Long term goals
 - Promote evidence-based treatment and quality improvement efforts for PTSD care
 - Develop a process to assess quality of psychosocial care
 - Develop measures that will be submitted for NQF endorsement

Project Timeline

March 1, 2012	TAG meeting
mid-March	Finalize concepts for measure development
March 2012 thru August 2012	Develop measure specifications and testing protocols
July/August 2012	TAG meeting (teleconference)
September/October 2012	Conduct focus groups to assess feasibility and usability
November/December 2012	Draft report summarizing project activities

Role of Technical Advisory Group

- Provide guidance on the prioritization of measure concepts
- Provide guidance on mechanisms and feasibility for developing measure concepts
- Review and provide feedback on the measure specifications
- Provide feedback and guidance on the testing plan

Goals of Today's Meeting

- Obtain guidance in identifying the most promising evidence-based measurement opportunities
- Obtain guidance on the types of measures that should be prioritized for development
- Obtain guidance on data sources for developing these measures

PTSD Overview

- PTSD is an anxiety disorder that some people develop following exposure to a traumatic event

Population	Prevalence
General Population	6.8%
Women	10%
Men	3.5%
Veterans	5% to 30%

- PTSD commonly co-occurs with depression, substance use, traumatic brain injury, and metabolic conditions

Evidence for PTSD Treatment and Care

Evidence for PTSD Care

- **Psychotherapy**
 - **Cognitive Behavior Therapy (CBT)**
 - Research and guidelines support the use of CBT
 - Most studies have examined exposure therapy (ET) and found it to be effective in reducing symptoms
 - **Eye Movement Desensitization and Reprocessing**
 - Different interpretations of research on effectiveness
- **Medication**
 - **SSRIs**
 - Conflicting interpretations of the research of effectiveness of SSRIs
 - Guidelines support use of SSRIs but vary in recommending them as first or second line of treatment

Evidence for PTSD Care

- **Care coordination**
 - Little research in PTSD; however, there is some evidence in the broader mental health field
- **Support services (housing, employment, peer support groups)**
 - Little research available in PTSD; some guidelines recommend assessing need for and/or provision of support services
- **Consumer experience**
 - Little research in PTSD, but evidence in broader mental health and health field supports measuring patient experience with care

Current PTSD Treatment and Care

PTSD Treatment Settings

- **In the general population, 57% of individuals with PTSD received mental health services within the past year**
 - 23% psychiatrist
 - 26% non-psychiatrist in specialty mental health
 - 31% general medical provider
 - 11% human services professional
 - 13% complimentary alternative medicine

Prevalence of PTSD Treatments Provided

- **Among patients with PTSD in an urban primary care setting**
 - 50% received mental health treatment
 - 35% received counseling and medication (SSRIs)
 - 29% received counseling only
 - 36% received medication only (SSRIs)
- **Among patients with PTSD in the VA**
 - 64% received mental health treatment
 - 54% received medication
 - 39% received counseling

Existing Related Measures

Existing PTSD Measures

Measure Name	Measure Developer
1. Number of evaluation and management visits with a prescribing provider following the start of a new treatment episode for patients undergoing pharmacotherapy	Greenberg et al.
2. Complicated PTSD with a new treatment episode of PTSD with no care by a licensed mental health provider	RAND/VA
3. Cognitive Behavior Therapy (CBT) for PTSD	RAND/VA
4. Proportion of patients with PTSD diagnosis who are monitored regarding symptom severity	RAND/VA
5. Proportion of patients with PTSD diagnosis who receive an adequate trial of selective serotonin reuptake inhibitors (SSRIs)	RAND/VA
6. Reassess severity of symptoms between the beginning of the second month and the end of the fourth month	RAND/VA
7. Reduction in target symptoms during the new treatment episode	RAND/VA

Existing PTSD Measures

Measure Name	Measure Developer
8. Comprehensive assessment: co-morbid psychiatric conditions, psychiatric history, and response to treatment	RAND/VA
9. All patients diagnosed with co-occurring SUD who are in a new treatment episode for COD should receive appropriate treatment for both their substance use disorder and mental health disorder	RAND/VA
10. Proportion of patients with co-occurring SUD and severe functional impairment that receive integrated SA and MH treatment	RAND/VA
11. Patients with a new treatment episode in specialty care receive baseline assessment of needs in the following domains: housing, social supports, employment	RAND/VA
12. Patients with identified need should be offered services for: social supports, housing, employment status	RAND/VA
13. Percentage of patient charts that document assessment for suicide ideation	RAND/VA
14. Family psychoeducation	Nelson & Wright 1996

MATHEMATICA
Policy Research

Measure Concepts & Points of Consideration

MATHEMATICA
Policy Research

NQF Framework

Importance

- Extent to which measure concept is evidence-based, has the potential for high impact on quality of care, and for which there is variation in or less than optimal performance

Scientific Acceptability

- Extent to which the measure properties produce consistent (reliable) and credible (valid) results about the quality of care

Usability

- Extent to which intended users can understand the results of the measure and find them useful for decision making and quality improvement

Feasibility

- Extent to which the data are readily available or can be collected without undue burden and can be implemented for performance measurement

Measure Concept Domains

- **Psychotherapy**
- **Pharmacotherapy**
- **Assessing/monitoring/treatment for other commonly occurring behavioral and physical health conditions**
- **Care coordination**
- **Patient experience**

Measurement Considerations

- **Level of specificity**
 - Is a treatment provided? vs. Were the appropriate components of a treatment provided?
- **Assessing psychotherapy**
 - Are there clearly defined treatment process elements?
 - Is there an established number of elements that must be provided in order for the treatment to be effective?
 - Are there feasible and reliable methods of measuring provision of treatment elements that don't involve direct observation?
 - Research on measuring CBT for depression via clinician and patient reported instruments could inform PTSD measure development

Measurement Considerations

- **Data Sources**
 - **Claims data**
 - Comparatively little burden for facilities to collect and report performance data using claims
 - Prior work at Mathematica suggests data on psychosocial treatment can not reliably be obtained in claims data
 - **Surveys**
 - Potential source for gathering information that may not be in charts
 - Feasibility of collecting and reporting survey data may concern providers

Measurement Considerations

- **Data Sources (continued)**
 - **Chart abstracted/Electronic health records**
 - Is provision of psychotherapy documented in a consistent way across providers?
 - What level of detail is provided (i.e., therapy vs. CBT vs. exposure therapy)?
 - How feasible is it to extract information on psychotherapy?
 - How feasible is it to measure fidelity to treatment implementation?

Measure Concept Domains

- **Psychotherapy**
- **Pharmacotherapy**
- **Assessing/monitoring/treatment for other commonly occurring behavioral and physical health conditions**
- **Care coordination**
- **Patient experience**

Questions

- **What measure domains and concepts should be prioritized?**
 - What measure concepts will advance quality of PTSD care?
 - What would be feasible to develop?
 - What level of specificity?

WRAP UP

- **Confirm final list of promising concepts**
- **Next TAG meeting (call) for July/August 2012**

APPENDIX B. TECHNICAL ADVISORY GROUP MEMBERS AND AFFILIATIONS

TABLE B.1. TAG Members and Affiliations	
Name	Affiliation
Francisca Azocar	Vice President, Research and Evaluation Behavioral Health Sciences OptumHealth Behavioral Solutions/United Behavioral Health
Julian Ford	Director of the Center for Trauma Response, Recovery, and Preparedness University of Connecticut Health Center Professor; University of Connecticut School of Medicine
Marcela Horvitz-Lennon	Physician Scientist; Adjunct Assistant Professor RAND Corporation
Lisa Jaycox	Senior Behavioral Scientist; Clinical Psychologist; Professor RAND Corporation
Stacey Kaltman	Associate Professor in the Department of Psychiatry and Assistant Director of the Center for Trauma and the Community Georgetown University Medical Center
Janice Krupnick	Director, Trauma and Loss Program Research Professor; Department of Psychiatry Georgetown University Medical Center
Dorene Loew	Program Director VA Palo Alto Health Care System Trauma Recovery Program
Linda Rosenberg	President & CEO National Council for Community Behavioral Healthcare
Dow Wieman	Senior Research Associate Human Services Research Institute
Daniel Williams	NAMI State Veteran Representative for Veterans council OIF Veteran U.S. Army

APPENDIX C. TECHNICAL EXPERT PANEL MEMBERS

TABLE C.1. TEP Members		
Name	Affiliation	Years of TEP Participation
Kathleen Chard, Ph.D.	Cincinnati VA Medical Center	2013-2014
Edna Foa, Ph.D. ^a	University of Pennsylvania	2012-2014
Patricia Resick, Ph.D., ABPP ^b	National Center for the Study of PTSD U.S. Department of Veteran Affairs	2012-2014
Barbara Rothbaum, Ph.D., ABPP	Emory University School of Medicine	2012-2014
Lori Zoellner, Ph.D.	University of Washington	2012-2014
a. Developer of prolonged exposure CBT; see Foa et al. (2007).		
b. Developer of CPT; see Resick and Schnicke (1996).		

APPENDIX D. LIST OF REVIEWED PTSD CLINICAL MANUALS

Beck, J.G., & Coffey, S.E. (2005). Group Cognitive Behavioral Treatment for PTSD: Treatment of Motor Vehicle Accident Survivors. *Cognitive and Behavioral Practice*, 12, 267-277

Blanchard, E.B., & Hickling, E. (2004). *After the crash: Psychological assessment and treatment of survivors of motor vehicle accidents* (2nd ed.). Washington, DC: American Psychological Association.

Foa, E.B., Hembree E., & Rothbaum, B.O. (2007). *Prolonged Exposure Therapy for PTSD: Emotional Processing of Traumatic Experiences, Therapist Guide*. New York, NY; Oxford University Press.

Resick, P.A., Monson, C.M., & Chard, K.M. (2010). *Cognitive processing therapy: Veteran/military version*. Washington, DC: U.S. Department of Veterans' Affairs.

Resick, P., & Schnicke, M. (1993). *Cognitive Processing Therapy for Rape Victims: A Treatment Manual* (Interpersonal Violence: The Practice Series).

Rothbaum, B.O., Foa, E.B., & Hembree, E. (2007). *Reclaiming Your Life from a Traumatic Experience: Client Workbook*. New York, NY: Oxford University Press.

Taylor, S., Thordarson, D.S., Maxfield, L., Fedoroff, I.C., Lovell, K. & Ogradnicsuk, J. (2003). Comparative Efficacy, Speed and Adverse Effect of Three PTSD Treatments: Exposure Therapy, EMDR and Relaxation Training: *Journal of Clinical and Consulting Psychology*, 71, (2), 330-338.

APPENDIX E. PARTICIPANT SURVEYS

Clinician Survey

SURVEY OF THE DELIVERY OF EVIDENCE BASED PSYCHOTHERAPY: CLINICIANS

Instructions: Please complete this survey based on the recent therapy session you had with _____ on _____.

Note: Not every therapeutic element will be delivered in every therapy session.

- Only endorse “yes” to those survey items that reflect the treatment you provided in this session. If the item did not occur in this session, please mark “no”.

Your responses will be kept confidential and will not be shared with your client, anyone within your organization, or anyone outside the Mathematica research team.

During this session on:	Please select one response		
	Yes	No	Don't remember
1. Did you set an agenda?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Did you go over the agenda with the client?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Did you provide background on the treatment rationales and concepts during this session (i.e., why you are asking the client to do something or explaining why something is occurring within the session)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Did you discuss or check-in on the client's treatment expectations (i.e., what will happen, how treatment will progress, expectations for improvement)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Did you and your client mutually set or check-in on goals for treatment?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Did you identify salient problem areas related to the trauma? - Problem areas might include self-blame, other blame, power and control issues, beliefs impacted by the trauma (e.g., the world is a dangerous place), self- esteem, safety, trust, intimacy, and perception of danger.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Did you use cognitive restructuring techniques (techniques to address cognitive issues such as negative thoughts, distortions, false beliefs or perceptions and replace them with accurate and more useful cognitions) to work on the identified problem areas?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

		Please select one response		
During this session on:		Yes	No	Don't remember
8.	<p>Did you use a Socratic discussion method, that is, statements or questions designed for the client to examine their beliefs?</p> <p>For example:</p> <ul style="list-style-type: none"> - How do you know this? Can you give me an example? - What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? - How did you come to this conclusion? What evidence do you have to justify this? 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9.	<p>Did you facilitate the development of alternative hypotheses (i.e., alternative viewpoints or explanations) to problematic beliefs?</p> <p>Examples of alternative hypotheses to problematic thinking might include:</p> <p>1.) <i>Distortion</i>: People in authority can't be trusted. More Helpful Thought: People in authority are individuals, and they don't all share the same strengths and weaknesses</p> <p>2.) <i>Distortion</i>: Everyone is out to hurt me. I can't trust anyone More Helpful Thought: There are some dangerous people out there, but not everyone is out to harm you</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10.	<p>Did you identify areas of trauma related avoidance, where the trauma has shifted or restricted daily patterns of living (i.e., the trauma has influenced daily functioning)? For example, a client may avoid places with loud noises and lots of people.</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11.	<p>Did you use techniques to systematically approach areas of trauma related avoidance, where the trauma has shifted daily patterns of living (i.e., the trauma has influenced daily functioning) from easier to more difficult situations?</p> <ul style="list-style-type: none"> - For example, a person in a motor vehicle accident may be fearful of driving. An approach from easy to more difficult might look like: <ul style="list-style-type: none"> o Easy: Encouraging the client to ride in a car as a passenger for a short period of time. o Difficult: Encouraging the client to drive on street and then a freeway, etc. 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

		Please select one response						
During this session on:		Yes	No	Don't remember				
12.	Did you use any of the following techniques to deal with trauma related avoidance?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
	a. Ask the client to imagine their traumatic experience for longer than 10 minutes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
	b. Ask the client to write about their traumatic experience	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
	c. Socratic discussion method (i.e., "How do you know this? Can you give me an example?")	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
	d. Real world experiments like visiting a place related to the traumatic experience with the client for longer than 10 minutes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
13.	Did you discuss and process the details of the client's recounting of the trauma, including the emotions surrounding the event?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
14.	Did you struggle to manage time for any of the reasons below:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
	- Client talked incessantly or tangentially							
	- Client had trouble keeping on task							
	- Session time was abbreviated							
	- You had trouble keeping the client on task							
15.	Were you directive (i.e., followed the agenda or guided the client to relevant discussion) during this session?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
16.	Did you ask your client for feedback or input on their treatment (i.e., "how is this working?"; "Are we working on things that you think are important?")? This would not include progress monitoring.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
17.	Did you ask your client for feedback on you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
18.	Did you assign your client homework or practice assignments (to be completed by the <i>next session</i>) to deal with issues surrounding PTSD symptoms (i.e., avoidance, thought monitoring, problematic beliefs, anxiety) or issues related to the trauma?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
19.	Did you review the assignment instructions and verify the client has a thorough understanding of the homework for the <i>next session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
20.	Did you address difficulties or barriers related to completing of homework from the <i>previous session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
21.	Did you work with your client to come up with solutions to difficulties, barriers, or issues in completing the homework from the <i>previous session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
22.	Did you review and discuss your client's homework assigned during the <i>previous session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
23.	When reviewing the homework assigned from the <i>previous session</i> , did you encourage the client or provide them with constructive feedback?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
		Never	Rarely	Occasionally	Sometimes	Often	Very Often	Always
24.	I am confident in my ability to help this client.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25.	I believe this client likes me as a therapist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

					Please select one response		
During this session on:					Yes	No	Don't remember
26.	This client and I have built mutual trust.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>For question part a., please think about the overall course of treatment with this client.</p> <p>For question part b., focus on care given during this selected treatment session.</p>							
					Yes	No	Don't remember
27.	a. Have you ever conducted a suicide risk assessment for this client?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	b. Did you conduct suicide risk assessment <i>during this session</i> ?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28.	a. Have you ever used information from your suicide risk assessment to influence treatment or monitor progress for this client?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	b. Did you use information from your suicide risk assessment to influence treatment or monitor progress <i>during this session</i> ?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
29.	a. Have you ever used any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	b. Did you use any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change <i>during this session</i> ?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30.	a. Have you ever provided the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.)?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	b. Did you provide the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.) <i>during this session</i> ?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31.	a. Have you ever provided the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma)?				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<ul style="list-style-type: none"> For example, this might include education on the nature of acquaintance rape vs. stranger rape, how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault. 						

During this session on:	Please select one response		
	Yes	No	Don't remember
b. Did you provide the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma) <i>during this session</i> ? <ul style="list-style-type: none"> For example, this might include education on the nature of acquaintance rape vs. stranger rape, or how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault. 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32. a. Have you ever provided the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Did you provide the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment) <i>during this session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Thank you for your participation!

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. Public reporting burden for this collection of information is estimated to average 20 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. **An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number.** Send comments regarding the burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to the Assistant Secretary of Planning and Evaluation, Room 415F, US Department of Health and Human Services, 200 Independence Ave. SW, Washington, DC 20201. Do not return the completed form to this address.

FORM APPROVED
 OMB NO. 0990-0418
 Exp. Date 05/31/2017

**Approved by NEIRB on 4/29/2014
 NEIRB Version No. 1.0**

Supervisor Survey

SURVEY OF THE DELIVERY OF EVIDENCE BASED PSYCHOTHERAPY: SUPERVISORS

Instructions: Please complete this survey based on your observation of the recent therapy session with _____ (clinician) and _____ (client).

Note: Not every therapeutic element will be delivered in every therapy session.

- Only endorse “yes” to those survey items that reflect the treatment you observed provided in this session.
- If the item did not occur in this session, please mark “no”.

Your responses will be kept confidential and not be shared with your supervisee, anyone within your organization, or anyone outside of the Mathematica research team.

1. What is the date of the session that you observed (i.e. date of session that supervision occurred)? _____ / _____ / _____
[mm/dd/yyyy]

During this session:	Please select one response:		
	Yes	No	Don't remember
2. Did the therapist set an agenda?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Did the therapist go over the agenda with the client?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Did the therapist provide background on the treatment rationales and concepts during this session (i.e., explaining why the therapist is asking the client to do something or explaining why something is occurring within the session)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Did the therapist discuss or check-in on the client's treatment expectations (i.e., what will happen, how treatment will progress, expectations for improvement)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Did the therapist and his/her client mutually set or check-in on goals for treatment?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Did the therapist identify salient problem areas related to the trauma? - Problem areas might include self-blame, other blame, power and control issues, beliefs impacted by the trauma (e.g., the world is a dangerous place), self- esteem, safety, trust, intimacy, and perception of danger.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Did the therapist use cognitive restructuring techniques (techniques to address cognitive issues such as negative thoughts, distortions, false beliefs or perceptions and replace them with accurate and more useful cognitions) to work on the identified problem areas?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

During this session:		Please select one response:		
		Yes	No	Don't remember
9.	<p>Did the therapist use a Socratic discussion method, that is, statements or questions designed for the client to examine his/her beliefs?</p> <p>For example:</p> <ul style="list-style-type: none"> - How do you know this? Can you give me an example? - What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? - How did you come to this conclusion? What evidence do you have to justify this? 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10.	<p>Did the therapist facilitate the development of alternative hypotheses (i.e., alternative viewpoints or explanations) to problematic beliefs?</p> <p>Examples of alternative hypotheses to problematic thinking might include:</p> <p>1.) <i>Distortion</i>: People in authority can't be trusted. More Helpful Thought: People in authority are individuals, and they don't all share the same strengths and weaknesses.</p> <p>2.) <i>Distortion</i>: Everyone is out to hurt me. I can't trust anyone. More Helpful Thought: There are some dangerous people out there, but not everyone is out to harm you .</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11.	<p>Did the therapist identify areas of trauma related avoidance, where the trauma has shifted or restricted daily patterns of living (i.e., the trauma has influenced daily functioning)? For example, a client may avoid places with loud noises and lots of people.</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12.	<p>Did the therapist use techniques to systematically approach areas of trauma related avoidance, where the trauma has shifted daily patterns of living (i.e., the trauma has influenced daily functioning) from easier to more difficult situations?</p> <ul style="list-style-type: none"> - For example, a person in a motor vehicle accident may be fearful of driving. An approach from easy to more difficult might look like: <ul style="list-style-type: none"> o Easy: Encouraging the client to ride in a car as a passenger for a short period of time. o Difficult: Encouraging the client to drive on street and then a freeway, etc. 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13.	<p>Did the therapist use any of the following techniques to deal with trauma related avoidance?</p> <p>a. Ask the client to imagine their traumatic experience for longer than 10 minutes</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

During this session:		Please select one response:						
		Yes	No	Don't remember				
	b. Ask the client to write about their traumatic experience	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
	c. Socratic discussion method (i.e., "How do you know this? Can you give me an example?")	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
	d. Real world experiments like visiting a place related to the traumatic experience with the client for longer than 10 minutes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
14.	Did the therapist discuss and process the details of the client's recounting of the trauma, including the emotions surrounding the event?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
15.	Did the therapist struggle to manage time for any of the reasons below: - Client talked incessantly or tangentially - Client had trouble keeping on task - Session time was abbreviated - The therapist had trouble keeping the client on task	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
16.	Was the therapist directive (i.e., followed the agenda or guided the client to relevant discussion) during this session?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
17.	Did the therapist ask the client for feedback or input on his/her treatment (i.e., "how is this working?"; "Are we working on things that you think are important?")? This would <i>exclude</i> progress monitoring.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
18.	Did the therapist ask the client for feedback on himself/herself?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
19.	Did the therapist assign his/her client homework or practice assignments (to be completed by the <i>next session</i>) to deal with issues surrounding PTSD symptoms (i.e., avoidance, thought monitoring, problematic beliefs, anxiety) or issues related to the trauma?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
20.	Did the therapist review the assignment instructions and verify the client has a thorough understanding of the homework for the <i>next session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
21.	Did the therapist address difficulties or barriers related to the completion of homework from the <i>previous session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
22.	Did the therapist work with the client to come up with solutions to difficulties, barriers, or issues in completing the homework from the <i>previous session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
23.	Did the therapist review and discuss the client's homework assigned during the <i>previous session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
24.	When reviewing the homework assigned from the <i>previous session</i> , did the therapist encourage the client or provide him/her with constructive feedback?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				
		Never	Rarely	Occasionally	Sometimes	Often	Very Often	Always
25.	I am confident in the therapist's ability to help this client.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

During this session:					Please select one response:		
					Yes	No	Don't remember
26.	I believe this client likes the therapist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
27.	This client and the therapist have built mutual trust.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>For question part a., please think about the overall course of treatment with this client.</p> <p>For question part b., focus on the care given during the selected session.</p>							
28.	a. To your knowledge, has the therapist ever conducted a suicide risk assessment with this client?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	b. Did the therapist conduct suicide risk assessment <i>during this session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
29.	a. To your knowledge, has the therapist ever used information from a suicide risk assessment to influence treatment or monitor progress with this client?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	b. Did the therapist use information from the suicide risk assessment to influence treatment or monitor progress <i>during this session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30.	a. To your knowledge, has the therapist used any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	b. Did the therapist use any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change <i>during this session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31.	a. To your knowledge has the therapist ever provided the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	b. Did the therapist provide the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.) <i>during this session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32.	a. To your knowledge, has the therapist ever provided the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<ul style="list-style-type: none"> For example, this might include education on the nature of acquaintance rape vs. stranger rape, how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault. 						
	b. Did the therapist provide the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma) <i>during this session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<ul style="list-style-type: none"> For example, this might include education on the nature of acquaintance rape vs. stranger rape, or how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault. 						

During this session:	Please select one response:		
	Yes	No	Don't remember
33. a. To your knowledge, has the therapist ever provided the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Did the therapist provide the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment) <i>during this session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Thank you for your participation!

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. Public reporting burden for this collection of information is estimated to average 20 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. **An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number.** Send comments regarding the burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to the Assistant Secretary of Planning and Evaluation, Room 415F, US Department of Health and Human Services, 200 Independence Ave. SW, Washington, DC 20201. Do not return the completed form to this address.

FORM APPROVED
OMB NO. 0990-0418
Exp. Date 05/31/2017

**Approved by NEIRB on 4/29/2014
NEIRB Version No. 1.0**

Client Survey

CLIENT SURVEY OF THE DELIVERY OF EVIDENCE BASED PSYCHOTHERAPY

Thank you for completing the Survey of the Delivery of Evidence Based Psychotherapy. Please read the following statement and choose “yes” or “no” below.

CONSENT TO PARTICIPATE IN THE SURVEY OF THE DELIVERY OF EVIDENCE BASED PSYCHOTHERAPY

I understand that:

I have been invited to take part in a survey that gathers information on the types of methods my therapist recently used to treat my PTSD.

The purpose of the study is to determine if the survey is a valid measure of quality of therapy for adults with PTSD.

My participation in this survey is voluntary, and I will not be penalized if I refuse to participate or decide to stop.

There is no cost to me to participate in the survey.

To the extent permitted by law, my survey responses will be kept private and secure.

- My information will only be used for this survey, and my name will not be associated with my answers.
- My individual answers will not be released to my therapist, the facility where I received treatment, or any other organization.
- Mathematica will summarize responses from all participants and share that information with ASPE, NIMH, or other organizations to make the survey better and to improve the quality of care for patients with PTSD.

In appreciation for completing the survey, I will receive a \$20 gift card from Mathematica.

I may change my mind and take back my permission at any time.

I can contact Melissa Azur, the project director, at mazur@mathematica-mpr.com or (202) 250-3518, or Kirsten Beronio, the Contract Officer Representative at ASPE, at Kirsten.Beronio@hhs.gov to get an answer about any questions I may have.

If I have questions about my rights as a research volunteer, or feel that I have been harmed in any way by participating in the study, I can call the New England Institutional Review Board, at 1-800-232-9570.

- **Please initial a response:**

_____ Yes, I consent to participate in the Survey of the Delivery of Evidence Based Psychotherapy.

Thank you, please continue to the next page to begin the survey.

_____ No, I do not consent to participate in the Survey of the Delivery of Evidence Based Psychotherapy.

Thank you for your response. If you have questions about the Survey of the Delivery of Evidence Based Psychotherapy or decide you would like to participate, please contact Melissa Azur, the project director, at mazur@mathematica-mpr.com or (202) 250-3518.

This survey is designed to understand and improve the quality of care provided to people with PTSD. Your thoughts on your current treatment are very important to us.

Please complete this survey based on the most recent session you had with your therapist. Not all of the below items will occur in every therapy session.

Choose “yes” only if the item occurred in the most recent therapy session.

Choose “no” if the item did not occur in the most recent therapy session.

If you cannot remember if an item did or did not occur, please choose “Don’t Remember”.

You may skip any question you do not feel comfortable answering.

Your responses will be kept confidential and will not be shared with your therapist or anyone outside the Mathematica research team.

		Please select one response:		
During this session:		Yes	No	Don't remember
1.	Did you and your therapist discuss an agenda or plan for your session?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	Did your therapist talk about or check-in on your expectations of how therapy will go?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	Did your therapist work with you to set goals you both agreed on?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	Did your therapist help you become aware of or realize feelings, views or thoughts in your life that have been influenced by your traumatic experience? <ul style="list-style-type: none"> • These might include feelings, views, or thoughts about being safe in the world, the presence of danger, trust, and self-esteem. 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	Did your therapist ask you several direct questions to <i>make you think critically about or examine</i> your thoughts, feelings, or beliefs? For example, your therapist might ask: <ul style="list-style-type: none"> • How do you know this? Can you give me an example? • What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? • How did you come to this conclusion? What evidence do you have to justify this? 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6.	Did your therapist offer other ways of thinking about your issues (e.g., problem areas or areas you want to work on) related to the trauma? For example: <ul style="list-style-type: none"> • Thought: “I can’t trust anyone.” • Thought suggested by therapist: “Some people can’t be trusted, but there are other people who are trustworthy.” 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7.	Did you and your therapist discuss people, events, or places you now avoid or stay away from because of your traumatic experience? For example, someone in a car accident might avoid driving on the freeway.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

		Please select one response:		
During this session:		Yes	No	Don't remember
8.	Did your therapist do <i>any</i> of the following things to help you deal with fear, anxiety or things you now avoid because of your trauma?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	a) Ask you to imagine or retell your traumatic experience for <i>longer than 10 minutes</i>			
	b) Ask you to write about your traumatic experience			
	c) Ask you questions to make you think critically about or examine your thoughts, feelings, or beliefs related to your fear, anxiety, and avoidance of things (i.e., "How do you know this? Can you give me an example?")			
	d) Ask you to do real world experiments like visiting a place related to the traumatic experience for <i>longer than 10 minutes</i>			
9.	After you described your traumatic experience, did you and your therapist discuss the details of what happened to you, how it impacted your life, or your emotions about the event?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10.	Did your therapist make good use of your session time today?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11.	Did your therapist ask for your opinion on how your treatment is going?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12.	Did your therapist ask for feedback on how she/he is doing in helping you recover from your PTSD?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13.	Did your therapist assign homework or practice <i>assignments (to be completed by the next session)</i> to work on your PTSD symptoms or problem areas?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14.	Did your therapist make sure you understood how to complete your homework for the <i>next session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15.	If you had problems completing your <i>previously</i> assigned homework, did your therapist work with you to come up with solutions to these problems?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16.	Did your therapist review and discuss your homework from the <i>previous session</i> ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17.	When reviewing the homework from the <i>previous session</i> , did your therapist encourage or provide you with constructive feedback?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

For the following questions, please think about the overall course of treatment with this therapist rather than the last session.

		Never	Rarely	Occasion-ally	Some-times	Often	Very Often	Always
18.	My therapist and I have built mutual trust.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19.	I am confident in my therapist's ability to help me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20.	I believe my therapist likes me as a person.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. a.	Has your therapist ever asked you if have had thoughts about committing suicide?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

During this session:		Please select one response:		
		Yes	No	Don't remember
b.	<i>During this session</i> , did your therapist ask you if you had thoughts about committing suicide?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. a.	Has your therapist ever asked you to answer questions about your PTSD symptoms? This might include completing a form before or after therapy.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b.	<i>During this session</i> , did your therapist ask you about your PTSD symptoms? This might include completing a form or survey before or after therapy.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. a.	Has your therapist ever provided information about PTSD and PTSD symptoms?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b.	<i>During this session</i> , did your therapist provide information about PTSD and PTSD symptoms?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. a.	Has your therapist ever provided with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? <ul style="list-style-type: none"> For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your viewpoints and beliefs 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b.	<i>During this session</i> , did your therapist ever provide you with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? <ul style="list-style-type: none"> For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your view points and beliefs 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. a.	Has your therapist ever explained how your particular treatment will work?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b.	<i>During this session</i> , did your therapist explain how your particular treatment will work?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In appreciation for completing this survey, we would like to mail you a \$20 gift card. If you would like to receive a gift card, please provide your mailing information below. We will only use this information to send you the gift card. You should receive your gift card in 2-3 weeks.

Name _____

Address _____

City _____

State _____

Zip _____

Phone _____

In late fall, we will be conducting telephone focus groups as a part of this study. The focus group will be about an hour long and participants will be paid a \$20 gift card after participating.

May we contact you about this?

Yes → Please provide best telephone number to reach you, if not provided above:

_____ - _____ - _____

No. _____

Thank you for completing this survey!

PLEASE MAIL TO MATHEMATICA POLICY RESEARCH IN THE PRE-PAID ENVELOPE PROVIDED.

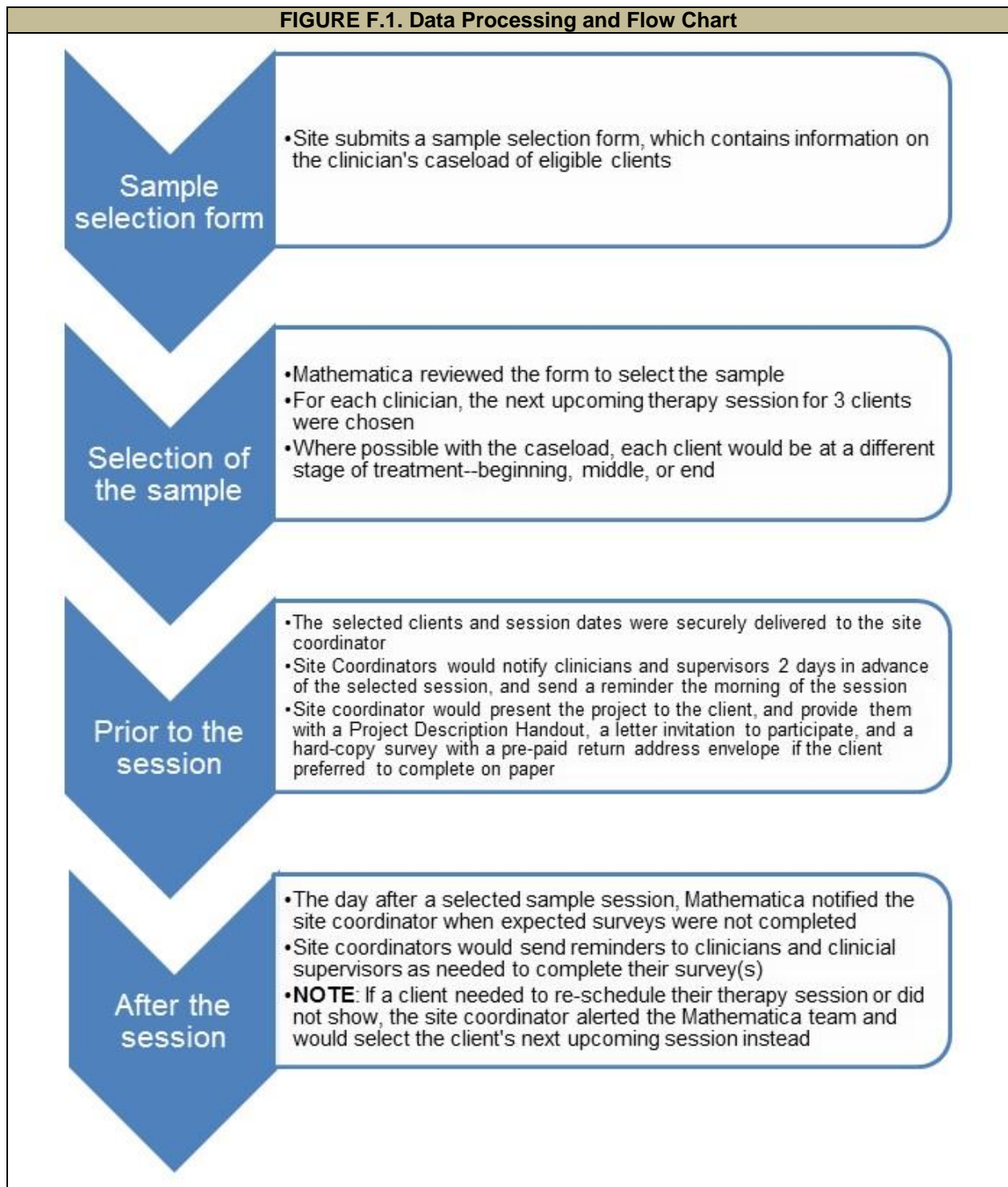
According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number. The valid OMB control number for this information collection is 0990-0418 . The time required to complete this information collection is estimated to average 10 minutes per response, including the time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. If you have comments concerning the accuracy of the time estimate(s) or suggestions for improving this form, please write to: U.S. Department of Health & Human Services, OS/OCIO/PRA, 200 Independence Ave., S.W., Suite 336-E, Washington D.C. 20201, Attention: PRA Reports

OMB NO. 0990-0418
Exp. Date 05/31/2017

**Approved by NEIRB on 4/29/2014
NEIRB Version No. 1.0**

APPENDIX F. DATA PROCESSING AND FLOW CHART

FIGURE F.1. Data Processing and Flow Chart



APPENDIX G. SUMMARY OF ITEM LEVEL RESPONSES

TABLE G.1. Summary of Item-Level Responses on the Clinician Survey				
Item	% Yes	% No	% Don't Remember	% Missing
Did you set an agenda?	60.42%	37.50%	1.04%	1.04%
Did you go over the agenda with the client?	54.17%	43.75%	2.08%	0.00%
Did you provide background on the treatment rationales and concepts during this session (i.e., why you are asking the client to do something or explaining why something is occurring within the session)?	77.08%	20.83%	2.08%	0.00%
Did you discuss or check-in on the client's treatment expectations (i.e., what will happen, how treatment will progress, expectations for improvement)?	58.33%	38.54%	3.13%	0.00%
Did you and your client mutually set or check-in on goals for treatment?	66.67%	26.04%	7.29%	0.00%
Did you identify salient problem areas related to the trauma? <ul style="list-style-type: none"> • Problem areas might include self-blame, other blame, power and control issues, beliefs impacted by the trauma (e.g., the world is a dangerous place), self-esteem, safety, trust, intimacy, and perception of danger. 	77.08%	21.88%	1.04%	0.00%
Did you use cognitive restructuring techniques (techniques to address cognitive issues such as negative thoughts, distortions, false beliefs or perceptions and replace them with accurate and more useful cognitions) to work on the identified problem areas?	63.54%	33.33%	3.13%	0.00%
Did you use a Socratic discussion method, that is, statements or questions designed for the client to examine their beliefs? For example: <ul style="list-style-type: none"> • How do you know this? Can you give me an example? • What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? • How did you come to this conclusion? What evidence do you have to justify this? 	62.50%	34.38%	3.13%	0.00%
Did you use any of the following techniques to deal with trauma related avoidance?				
Did you facilitate the development of alternative hypotheses (i.e., alternative viewpoints or explanations) to problematic beliefs? Examples of alternative hypotheses to problematic thinking might include: <ol style="list-style-type: none"> 1) Distortion: People in authority can't be trusted. More Helpful Thought: People in authority are individuals, and they don't all share the same strengths and weaknesses 2) Distortion: Everyone is out to hurt me. I can't trust anyone. More Helpful Thought: There are some dangerous people out there, but not everyone is out to harm you 	57.29%	37.50%	5.21%	0.00%
Did you identify areas of trauma related avoidance, where the trauma has shifted or restricted daily patterns of living (i.e., the trauma has influenced daily functioning)? For example, a client may avoid places with loud noises and lots of people.	59.38%	39.58%	1.04%	0.00%

TABLE G.1 (continued)				
Item	% Yes	% No	% Don't Remember	% Missing
Did you use techniques to systematically approach areas of trauma related avoidance, where the trauma has shifted daily patterns of living (i.e., the trauma has influenced daily functioning) from easier to more difficult situations? - For example, a person in a motor vehicle accident may be fearful of driving. An approach from easy to more difficult might look like: <ul style="list-style-type: none"> • Easy: Encouraging the client to ride in a car as a passenger for a short period of time. • Difficult: Encouraging the client to drive on street and then a freeway, etc. 	35.42%	63.54%	1.04%	0.00%
a. Ask the client to imagine their traumatic experience for longer than 10 minutes	14.58%	84.38%	1.04%	0.00%
b. Ask the client to write about their traumatic experience	21.88%	75.00%	3.13%	0.00%
c. Socratic discussion method (i.e., How do you know this? Can you give me an example?)	46.88%	50.00%	2.08%	1.04%
d. Real world experiments like visiting a place related to the traumatic experience with the client for longer than 10 minutes	10.42%	86.46%	3.13%	0.00%
Did you discuss and process the details of the client's recounting of the trauma, including the emotions surrounding the event?	45.83%	53.13%	1.04%	0.00%
Did you struggle to manage time for any of the reasons below: <ul style="list-style-type: none"> • Client talked incessantly or tangentially • Client had trouble keeping on task • Session time was abbreviated • You had trouble keeping the Client on task 	31.25%	66.67%	2.08%	0.00%
Were you directive (i.e., followed the agenda or guided the client to relevant discussion) during this session?	88.54%	10.42%	1.04%	0.00%
Did you ask your client for feedback or input on their treatment (i.e., how is this working?; Are we working on things that you think are important?); This would not include progress monitoring.	65.63%	30.21%	4.17%	0.00%
Did you ask your client for feedback on you?	29.17%	67.71%	3.13%	0.00%
Did you assign your client homework or practice assignments (to be completed by the next session) to deal with issues surrounding PTSD symptoms (i.e., avoidance, thought monitoring, problematic beliefs, anxiety) or issues related to the trauma?	63.54%	33.33%	3.13%	0.00%
Did you review the assignment instructions and verify the client has a thorough understanding of the homework for the next session?	47.92%	41.67%	9.38%	1.04%
Did you address difficulties or barriers related to completing of homework from the previous session?	28.13%	65.63%	6.25%	0.00%
Did you work with your client to come up with solutions to difficulties, barriers, or issues in completing the homework from the previous session?	31.25%	64.58%	4.17%	0.00%
Did you review and discuss your client's homework assigned during the previous session?	31.25%	63.54%	5.21%	0.00%
For question part a., please think about the overall course of treatment with this client. For question part b., focus on care given during this selecte+A32:A42d treatment				
When reviewing the homework assigned from the previous session, did you encourage the client or provide them with constructive feedback?	30.21%	60.42%	7.29%	2.08%
a. Have you ever conducted a suicide risk assessment for this client?	83.33%	14.58%	1.04%	1.04%
b. Did you conduct suicide risk assessment during this session?	25.00%	70.83%	4.17%	0.00%
a. Have you ever used information from your suicide risk assessment to influence treatment or monitor progress for this client?	63.54%	35.42%	1.04%	0.00%

TABLE G.1 (continued)				
Item	% Yes	% No	% Don't Remember	% Missing
b. Did you use information from your suicide risk assessment to influence treatment or monitor progress during this session?	27.08%	70.83%	2.08%	0.00%
a. Have you ever used any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change?	35.42%	64.58%	0.00%	0.00%
b. Did you use any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change during this session?	9.38%	89.58%	0.00%	1.04%
a. Have you ever provided the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.)?	92.71%	6.25%	1.04%	0.00%
b. Did you provide the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.) during this session?	55.21%	42.71%	2.08%	0.00%
a. Have you ever provided the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma)? • For example, this might include education on the nature of acquaintance rape vs. stranger rape, how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault.	67.71%	28.13%	3.13%	1.04%
b. Did you provide the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma) during this session? • For example, this might include education on the nature of acquaintance rape vs. stranger rape, or how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault.	31.25%	64.58%	4.17%	0.00%
a. Have you ever provided the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment)?	82.29%	12.50%	3.13%	2.08%
b. Did you provide the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment) during this session?	36.46%	62.50%	1.04%	0.00%
	Mean (Range)	Standard Deviation	% Don't Remember	% Missing
I am confident in my ability to help this client.	5.53 (3-7)	0.98	NA*	0.00%
I believe this client likes me as a therapist.	5.89 (4-7)	0.84	NA*	0.00%
This client and I have built mutual trust.	5.97 (3-7)	0.90	NA*	0.00%

* The 7-point Likert scale questions did not allow a "Don't Remember" option.

TABLE G.2. Summary of Item-Level Responses on the Supervisor Survey				
Item	% Yes	% No	% Don't Remember	% Missing
Did you use any of the following techniques to deal with trauma related avoidance?				
Did the therapist set an agenda?	50.00%	33.67%	15.31%	1.02%
Did the therapist go over the agenda with the client?	44.90%	35.71%	18.37%	1.02%
Did the therapist provide background on the treatment rationales and concepts during this session (i.e., why you are asking the client to do something or explaining why something is occurring within the session)?	52.04%	24.49%	22.45%	1.02%
Did the therapist discuss or check-in on the client's treatment expectations (i.e., what will happen, how treatment will progress, expectations for improvement)?	47.96%	28.57%	22.45%	1.02%
Did the therapist and his/her client mutually set or check-in on goals for treatment?	58.16%	22.45%	17.35%	2.04%
Did the therapist identify salient problem areas related to the trauma? <ul style="list-style-type: none"> • Problem areas might include self-blame, other blame, power and control issues, beliefs impacted by the trauma (e.g., the world is a dangerous place), self-esteem, safety, trust, intimacy, and perception of danger. 	74.49%	22.45%	2.04%	1.02%
Did the therapist use cognitive restructuring techniques (techniques to address cognitive issues such as negative thoughts, distortions, false beliefs or perceptions and replace them with accurate and more useful cognitions) to work on the identified problem areas?	70.41%	21.43%	7.14%	1.02%
Did the therapist use a Socratic discussion method, that is, statements or questions designed for the client to examine their beliefs? <ul style="list-style-type: none"> • How do you know this? Can you give me an example? • What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? • How did you come to this conclusion? What evidence do you have to justify this? 	67.35%	16.33%	15.31%	1.02%
Did the therapist facilitate the development of alternative hypotheses (i.e., alternative viewpoints or explanations) to problematic beliefs? Examples of alternative hypotheses to problematic thinking might include: 1) Distortion: People in authority can't be trusted. (More Helpful Thought: People in authority are individuals, and they don't all share the same strengths and weaknesses) 2) Distortion: Everyone is out to hurt me. I can't trust anyone (More Helpful Thought: There are some dangerous people out there, but not everyone is out to harm you)	61.22%	17.35%	20.41%	1.02%
Did the therapist identify areas of trauma related avoidance, where the trauma has shifted or restricted daily patterns of living (i.e., the trauma has influenced daily functioning)? For example, a client may avoid places with loud noises and lots of people.	50.00%	36.73%	11.22%	2.04%
Did the therapist use techniques to systematically approach areas of trauma related avoidance, where the trauma has shifted daily patterns of living (i.e., the trauma has influenced daily functioning) from easier to more difficult situations? For example, a person in a motor vehicle accident may be fearful of driving. An approach from easy to more difficult might look like: <ul style="list-style-type: none"> • Easy: Encouraging the client to ride in a car as a passenger for a short period of time. • Difficult: Encouraging the client to drive on street and then a freeway, etc. 	33.67%	59.18%	6.12%	1.02%
a. Ask the client to imagine their traumatic experience for longer than 10 minutes	11.22%	80.61%	8.16%	0.00%

TABLE G.2 (continued)				
Item	% Yes	% No	% Don't Remember	% Missing
b. Ask the client to write about their traumatic experience	19.39%	71.43%	7.14%	2.04%
c. Socratic discussion method (i.e., How do you know this? Can you give me an example?)	47.96%	42.86%	6.12%	3.06%
d. Real world experiments like visiting a place related to the traumatic experience with the client for longer than 10 minutes	14.29%	78.57%	5.10%	2.04%
Did the therapist discuss and process the details of the client's recounting of the trauma, including the emotions surrounding the event?	41.84%	52.04%	4.08%	2.04%
Did the therapist struggle to manage time for any of the reasons below: <ul style="list-style-type: none"> • Client talked incessantly or tangentially • Client had trouble keeping on task • Session time was abbreviated • You had trouble keeping the Client on task 	13.27%	71.43%	13.27%	2.04%
Was the therapist directive (i.e., followed the agenda or guided the client to relevant discussion) during this session?	79.59%	10.20%	10.20%	0.00%
Did the therapist ask his/her client for feedback or input on their treatment (i.e., how is this working?; Are we working on things that you think are important?); This would not include progress monitoring.	57.14%	27.55%	13.27%	2.04%
Did the therapist ask your client for feedback on himself/herself?	26.53%	55.10%	18.37%	0.00%
Did the therapist assign his/her client homework or practice assignments (to be completed by the next session) to deal with issues surrounding PTSD symptoms (i.e., avoidance, thought monitoring, problematic beliefs, anxiety) or issues related to the trauma?	48.98%	47.96%	3.06%	0.00%
Did the therapist review the assignment instructions and verify the client has a thorough understanding of the homework for the next session?	39.80%	45.92%	12.24%	2.04%
Did the therapist address difficulties or barriers related to completing of homework from the previous session?	18.37%	72.45%	8.16%	1.02%
Did the therapist work with his/her client to come up with solutions to difficulties, barriers, or issues in completing the homework from the previous session?	14.29%	75.51%	9.18%	1.02%
Did the therapist review and discuss the client's homework assigned during the previous session?	22.45%	64.29%	11.22%	2.04%
When reviewing the homework assigned from the previous session, did the therapist encourage the client or provide him/her with constructive feedback?	15.31%	68.37%	16.33%	0.00%
For question part a., please think about the overall course of treatment with this client.				
For question part b., focus on caregiving given during this selected treatment session.				
a. To your knowledge, has the therapist ever conducted a suicide risk assessment with this client?	88.78%	7.14%	2.04%	2.04%
b. Did the therapist conduct a suicide risk assessment during this session?	23.47%	71.43%	5.10%	0.00%
a. To your knowledge, has the therapist ever used information from your suicide risk assessment to influence treatment or monitor progress with this client?	71.43%	17.35%	11.22%	0.00%
b. Did the therapist use information from the suicide risk assessment to influence treatment or monitor progress during this session?	18.37%	72.45%	8.16%	1.02%
a. To your knowledge, has the therapist used any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change?	28.57%	61.22%	10.20%	0.00%
b. Did the therapist use any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change during this session?	7.14%	84.69%	8.16%	0.00%

TABLE G.2 (continued)				
Item	% Yes	% No	% Don't Remember	% Missing
a. To your knowledge has the therapist ever provided the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.)?	83.67%	2.04%	13.27%	1.02%
b. Did the therapist provide the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.) during this session?	61.22%	32.65%	6.12%	0.00%
a. To your knowledge, has the therapist ever provided the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma)? • For example, this might include education on the nature of acquaintance rape vs. stranger rape, how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault.	64.29%	14.29%	20.41%	1.02%
b. Did the therapist provide the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma) during this session? • For example, this might include education on the nature of acquaintance rape vs. stranger rape, or how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault.	39.80%	52.04%	7.14%	1.02%
a. To your knowledge, has the therapist ever provided the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment)?	65.31%	13.27%	20.41%	1.02%
b. Did the therapist provide the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment) during this session?	33.67%	55.10%	10.20%	1.02%
	Mean (Range)	Standard Deviation	% Don't Remember	% Missing
I am confident in the therapist's ability to help this client.	6.38 (3-7)	0.82	NA	0.00%
I believe this client likes the therapist.	6.18 (3-7)	0.74	NA	0.00%
This client and the therapist have built mutual trust.	6.23 (3-7)	0.74	NA	0.00%

* The 7-point Likert scale questions did not allow a "Don't Remember" option.

TABLE G.3. Summary of Item-Level Responses on the Client Survey				
Question	% Yes	% No	% Don't Remember	% Missing
Did you and your therapist discuss an agenda or plan for your session?	89.74%	10.26%	0.00%	0.00%
Did your therapist talk about or check-in on your expectations of how therapy will go?	79.49%	15.38%	5.13%	0.00%
Did your therapist work with you to set goals you both agreed on?	91.03%	5.13%	0.00%	1.28%
Did your therapist help you become aware of or realize feelings, views or thoughts in your life that have been influenced by your traumatic experience? <ul style="list-style-type: none"> • These might include feelings, views, or thoughts about being safe in the world, the presence of danger, trust, and self-esteem. 	92.31%	7.69%	0.00%	0.00%
Did your therapist ask you several direct questions to make you think critically about or examine your thoughts, feelings, or beliefs? For example, your therapist might ask: <ul style="list-style-type: none"> • How do you know this? Can you give me an example? • What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? • How did you come to this conclusion? What evidence do you have to justify this? 	89.74%	7.69%	0.00%	0.00%
Did your therapist offer other ways of thinking about your issues (e.g., problem areas or areas you want to work on) related to the trauma? For example: <ul style="list-style-type: none"> • Thought: "I can't trust anyone." • Thought suggested by therapist: "Some people can't be trusted, but there are other people who are trustworthy." 	91.03%	5.13%	3.85%	0.00%
Did you and your therapist discuss people, events, or places you now avoid or stay away from because of your traumatic experience? For example, someone in a car accident might avoid driving on the freeway.	82.05%	11.54%	6.41%	0.00%
Did your therapist do any of the following things to help you deal with fear, anxiety or things you now avoid because of your trauma? <ol style="list-style-type: none"> Ask you to imagine or retell your traumatic experience for longer than 10 minutes Ask you to write about your traumatic experience Ask you questions to make you think critically about or examine your thoughts, feelings, or beliefs related to your fear, anxiety, and avoidance of things (i.e., "How do you know this? Can you give me an example?") Ask you to do real world experiments like visiting a place related to the traumatic experience for longer than 10 minutes 	70.51%	23.08%	2.56%	3.85%
After you described your traumatic experience, did you and your therapist discuss the details of what happened to you, how it impacted your life, or your emotions about the event?	83.33%	14.10%	2.56%	0.00%
Did your therapist make good use of your session time today?	97.44%	1.28%	0.00%	1.28%
Did your therapist ask for your opinion on how your treatment is going?	75.64%	19.23%	5.13%	0.00%
Did your therapist ask for feedback on how she/he is doing in helping you recover from your PTSD?	47.44%	35.90%	11.54%	5.13%
Did your therapist assign homework or practice assignments (to be completed by the next session) to work on your PTSD symptoms or problem areas?	60.26%	35.90%	2.56%	1.28%
Did your therapist make sure you understood how to complete your homework for the next session?	65.38%	29.49%	3.85%	1.28%

TABLE G.3 (continue)				
Question	% Yes	% No	% Don't Remember	% Missing
If you had problems completing your previously assigned homework, did your therapist work with you to come up with solutions to these problems?	62.82%	24.36%	8.97%	3.85%
Did your therapist review and discuss your homework from the previous session?	56.41%	29.49%	7.69%	6.41%
When reviewing the homework from the previous session, did your therapist encourage or provide you with constructive feedback?	57.69%	25.64%	11.54%	5.13%
a. Has your therapist ever asked you if have had thoughts about committing suicide?	88.46%	8.97%	1.28%	1.28%
b. During this session, did your therapist ask you if you had thoughts about committing suicide?	41.03%	52.56%	3.85%	2.56%
a. Has your therapist ever asked you to answer questions about your PTSD symptoms? This might include completing a form before or after therapy.	73.08%	15.38%	7.69%	3.85%
b. During this session, did your therapist ask you about your PTSD symptoms? This might include completing a form or survey before or after therapy.	65.38%	26.92%	5.13%	2.56%
a. Has your therapist ever provided information about PTSD and PTSD symptoms?	74.36%	15.38%	7.69%	2.56%
b. During this session, did your therapist provide information about PTSD and PTSD symptoms?	50.00%	41.03%	6.41%	2.56%
a. Has your therapist ever provided with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? • For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your viewpoints and beliefs	70.51%	19.23%	8.97%	1.28%
b. During this session, did your therapist ever provide you with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? • For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your view points and beliefs	52.56%	37.18%	7.69%	2.56%
a. Has your therapist ever explained how your particular treatment will work?	92.31%	3.85%	0.00%	3.85%
b. During this session, did your therapist explain how your particular treatment will work?	73.08%	16.67%	7.69%	2.56%
	Mean (Range)	Standard Deviation	% Don't Remember	% Missing
My therapist and I have built mutual trust.	6.44 (2-7)	0.98	NA	1.28%
I am confident in my therapist's ability to help me.	6.29 (2-7)	1.17	NA	1.28%
I believe my therapist likes me as a person.	6.25 (2-7)	1.16	NA	1.28%

APPENDIX H. EXPLORATORY AND CONFIRMATORY FACTOR ANALYSIS MODEL FIT STATISTICS

TABLE H.1. EFA Model Fit Statistics				
Model Fit Information	EFA Solution			
	5-Factor	6-Factor	7-Factor	8-Factor
RMSEA	0.043	0.037	0.032	0.027
CFI	0.958	0.971	0.979	0.986
TLI	0.946	0.960	0.969	0.979
Number of parameters	203	240	276	311
Degrees of freedom	661	624	588	553

NOTES: RMSEA is the Root Mean Squared Error of Approximation (<0.1 is the minimally accepted cut-off; <0.05 is desired); CFI is Comparative Fit Index (>0.9 is the minimally accepted cut-off; >0.95 is desired); TLI is a Tucker-Lewis Index (>0.9 is the minimally accepted cut-off; >0.95 is desired).

TABLE H.2. CFAA Model Fit Statistics				
Estimator	Fit Statistics	Sample		
		Supervisor	Clinician	Client
Bayesian	PPP	0.180	0.148	0.370
WLSMV	RMSEA	0.053	0.054	0.054
	CFI	0.924	0.808	0.893
	TLI	0.924	0.797	0.883

NOTE: PPP is a Posterior Predictive P-value; RMSEA is the Root Mean Squared Error of Approximation (<0.1 is the minimally accepted cut-off; <0.05 is desired); CFI is Comparative Fit Index (>0.9 is the minimally accepted cut-off; >0.95 is desired); TLI is a Tucker-Lewis Index (>0.9 is the minimally accepted cut-off; >0.95 is desired).

APPENDIX I. CONFIRMATORY FACTOR ANALYSIS FINAL MODEL SOLUTION

TABLE I.1. CFA Final Model Solution					
Supervisor Sample		Clinicians Sample		Clients Sample	
Item and Number	λ	Item	λ	Item	λ
FACTOR 1 “Structuring and Conducting the Session”					
2. AGENDA	0.961	1. AGENDA	0.975	1. AGENDA	0.775
5. EXPECTATIONS	0.854	4. EXPECTATIONS	0.604	2. EXPECTATIONS	0.896
6. GOALS	0.746	5. GOALS	0.414	3. GOALS	0.806
17. TX FEEDBACK	0.937	16. TX FEEDBACK	0.487	11. TX FEEDBACK	0.777
33.B. TODAY OUTLINE	0.799	32.B. TODAY OUTLINE	0.829	25.B. TODAY OUTLINE	0.870
30.B. TODAY INSTRU	NA	29.B. TODAY INSTRU	0.843	22.B. TODAY INSTRU	0.680
33.A. EVER OUTLINE	NA	32.A. EVER OUTLINE	0.440	25.A. EVER OUTLINE	0.787
3. REVIEW AGENDA	0.982	2. REVIEW AGENDA	0.973	QUESTION NOT INCLUDED IN SURVEY	NA*
16. DIRECTIVE	0.759	15. DIRECTIVE	0.801	QUESTION NOT INCLUDED IN SURVEY	NA*
FACTOR 2: “Psychoeducation and Therapeutic Techniques”					
14. DISCUSS	0.799	13. DISCUSS	0.636	9. DISCUSS	0.932
31.B. TODAY SYMP EDU	0.854	30.B. TODAY SYMP EDU	0.654	23.B. TODAY SYMP EDU	0.926
31.A. EVER SYMP EDU	0.679	30.A. EVER SYMP EDU	0.896	23.A. EVER SYMP EDU	0.679
32.B. TODAY TRAUMA ED	0.930	31.B. TODAY TRAUMA ED	0.490	24.B. TODAY TRAUMA ED	0.943
13. OVERALL TECHNIQUES	0.942	12. OVERALL TECHNIQUES	0.978	8. OVERALL TECHNIQUES	0.721
32.A. EVER TRAUMA ED	0.829	31.A. EVER TRAUMA ED	NA	24.A. EVER TRAUM ED	0.830
7. IDENTIFY	0.829	6. IDENTIFY	0.654	4. IDENTIFY	0.635
11. OTHER IDENTIFY	0.846	10. OTHER IDENTIFY	0.516	7. OTHER IDENTIFY	0.452
13.A. IMAGINE	0.731	12.A. IMAGINE	0.685	QUESTION NOT INCLUDED IN SURVEY	NA*
13.B. WRITE	0.763	12.B. WRITE	0.875	QUESTION NOT INCLUDED IN SURVEY	NA*
13.C. OTHER SOCRAT	0.913	12.C. OTHER SOCRAT	0.929	QUESTION NOT INCLUDED IN SURVEY	NA*
13.D. REAL	0.704	12.D. REAL	0.742	QUESTION NOT INCLUDED IN SURVEY	NA*
4. BACKGROUND	0.811	3. BACKGROUND	0.678	QUESTION NOT INCLUDED IN SURVEY	NA*
8. COG RESTRUC	0.555	7. COG RESTRUC	0.582	QUESTION NOT INCLUDED IN SURVEY	NA*
9. SOCRAT	0.470	8. SOCRAT	0.788	5. SOCRAT	NA

TABLE I.1 (continued)					
Supervisor Sample		Clinicians Sample		Clients Sample	
Item and Number	λ	Item	λ	Item	λ
12. TECHNIQUES	0.790	11. TECHNIQUES	0.517	QUESTION NOT INCLUDED IN SURVEY	NA*
19. ASSIGN	0.590	18. ASSIGN	NA	13. ASSIGN	NA
20. REVIEW INSTRUC	0.518	19. REVIEW INSTRUC	NA	14. REVIEW INSTRUC	NA
FACTOR 3: "Therapeutic Alliance"					
24. CONFIDENT	0.662	25. CONFIDENT	0.653	19. CONFIDENT	0.776
25. LIKES	0.849	26. LIKES	0.826	20. LIKES	0.772
26. TRUST	0.950	27. TRUST	0.867	18. TRUST	0.901
FACTOR 4: "Suicide Assessment"					
28.b. TODAY SUIC	0.962	27.b. TODAY SUIC	0.762	21.b.TODAY SUIC	0.630
29.b. TODAY USE SUIC	0.962	28.b.TODAY USE SUIC	0.920	21.b. TODAY USE SUIC	NA
10. FACILITATE	NA	9. FACILITATE	0.484	6. FACILITATE	0.848
18. TH FEEDBACK	NA	17. TH FEEDBACK	0.675	12. TH FEEDBACK	0.918
15. STRUGGLE	0.555	14. STRUGGLE	NA	10. STRUGGLE	0.789
29.a. EVER USE SUIC	NA	28.a. EVER USE SUIC	0.578	QUESTION NOT INCLUDED IN SURVEY	NA*
9. SOCRAT	NA	8. SOCRAT	NA	5. SOCRAT	0.584
FACTOR 5: "Homework"					
19. ASSIGN	0.751	18. ASSIGN	0.963	13. ASSIGN	0.969
20. REVIEW INSTRUC	0.801	19. REVIEW INSTRUC	0.940	14. REVIEW INSTRUC	0.967
22. SOLUTION	0.841	21. SOLUTION	0.795	15. SOLUTION	0.794
23. REVIEW HMMWK	0.942	22. REVIEW HMMWK	0.826	16. REVIEW HMMWK	0.879
21. ADDRESS	0.935	20. ADDRESS	0.914	QUESTION NOT INCLUDED IN SURVEY	NA*
NA = Question did not load onto this factor for the model indicated in the column header. NA* = Question was not included in the client version of the survey.					

APPENDIX J. SUMMARY OF THE RELIABILITY ANALYSES

TABLE J.1. Summary of the Reliability Analysis for Factor 1: Structuring and Conducting the Session				
Item	N	Item Difficulty	Item Variance	Item-Rest Correlate
Supervisor Sample				
2. agenda	66	0.62	0.24	0.77
5. expectations	66	0.61	0.24	0.69
6. goals	66	0.70	0.21	0.53
18. tx feedback	66	0.70	0.21	0.76
33.b.today outline	66	0.45	0.25	0.53
3. review agenda	66	0.59	0.24	0.81
16. directive	66	0.86	0.12	0.54
31.b. todayinstru	NA	NA	NA	NA
34.a. ever outline	NA	NA	NA	NA
Clinician Sample				
1. agenda	80	0.56	0.25	0.63
4..expectations	80	0.58	0.24	0.49
5. goals	80	0.70	0.21	0.31
17. tx feedback	80	0.69	0.21	0.44
32.b.today outline	80	0.36	0.23	0.63
2. review agenda	80	0.51	0.25	0.65
15. directive	80	0.88	0.11	0.35
30.b. todayinstru	80	0.09	0.08	0.35
33.a. ever outline	80	0.86	0.12	0.27
Client Sample				
1. agenda	57	0.89	0.09	0.41
2. expectations	57	0.86	0.12	0.67
3. goals	57	0.95	0.05	0.42
11. tx feedback	57	0.77	0.18	0.44
25.b. today outline	57	0.82	0.14	0.60
Question not on client survey	NA*	NA*	NA*	NA*
Question not on client survey	NA*	NA*	NA*	NA*
22.b. today instru	57	0.68	0.22	0.51
25.a. ever outline	57	0.95	0.05	0.37
NA = Question did not load onto this factor for the model indicated in the column header. NA* = Question was not included in the client version of the survey.				

**TABLE J.2. Summary of the Reliability Analysis for Factor 2:
Psychoeducation and Therapeutic Techniques**

Item	N	Item Difficulty	Item Variance	Item-Rest Correlate
Supervisor Sample				
13.a. imagine	49	0.16	0.14	0.34
12.b. write	49	0.29	0.20	0.46
13.c. other socrat	49	0.69	0.21	0.70
13.d. real	49	0.12	0.11	0.33
4. background	49	0.71	0.20	0.52
7. identify	49	0.80	0.16	0.53
8. cog restruc	49	0.82	0.15	0.08
9. socrat	49	0.76	0.18	0.52
11. Other identify	49	0.57	0.24	0.64
12. techniques	49	0.37	0.23	0.48
14. discuss	49	0.47	0.25	0.47
19. assign	49	0.45	0.25	0.60
20. Review instruc	49	0.45	0.25	0.48
31.a. ever symp edu	NA	NA	NA	NA
31.b. today symp edu	49	0.73	0.19	0.61
32.a. Ever trauma ed	49	0.82	0.15	0.58
32.b. today trauma ed	49	0.57	0.24	0.75
12. Overall techniques	49	0.73	0.19	0.72
Clinician Sample				
12.a. imagine	80	0.16	0.14	0.44
13.write	80	0.18	0.14	0.52
12.c. other socrat	80	0.46	0.25	0.61
12.d. real	80	0.13	0.11	0.39
3. background	80	0.79	0.17	0.40
6. identify	80	0.75	0.19	0.49
7. cog restruc	80	0.65	0.23	0.36
8. socrat	80	0.61	0.24	0.46
10. other identify	80	0.60	0.24	0.31
11. techniques	80	0.36	0.23	0.37
13. discuss	80	0.45	0.25	0.43
18. assign	NA	NA	NA	NA
19. review instruc	NA	NA	NA	NA
30.a. Ever symp edu	80	0.94	0.06	0.41
30.b. today symp edu	80	0.59	0.24	0.46
31.a. Ever trauma ed	NA	NA	NA	NA
31.b. Today trauma ed	80	0.28	0.20	0.38
12. Overall techniques	80	0.54	0.25	0.66
Client Sample				
Question not included in client survey	NA*	NA*	NA*	NA*
Question not included in client survey	NA*	NA*	NA*	NA*
Question not included in client survey	NA*	NA*	NA*	NA*
Question not included in client survey	NA*	NA*	NA*	NA*
Question not included in client survey	NA*	NA*	NA*	NA*
4. identify	51	0.92	0.07	0.23
Question not included in client survey	NA*	NA*	NA*	NA*
5. socrat	NA	NA	NA	NA
7. other identify	51	0.90	0.09	0.05
Question not included in client survey	NA*	NA*	NA*	NA*
9. discuss	51	0.88	0.10	0.61
13. assign	NA	NA	NA	NA
14. review instruc	NA	NA	NA	NA
23.a. ever symp edu	51	0.84	0.13	0.48
23.b. today symp edu	51	0.63	0.23	0.65
24.a. ever trauma ed	51	0.78	0.17	0.53
24.b. today trauma ed	51	0.63	0.23	0.62
8. Overall techniques	51	0.78	0.17	0.43
NA = Question did not load onto this factor for the model indicated in the column header. NA* = Question was not included in the client version of the survey.				

TABLE J.3. Summary of the Reliability Analysis for Factor 3, Factor 4, and Factor 5				
Factor 3: Therapeutic Alliance				
Item	Obs	Item-Domain Correlation	Inter-Item Covariance	Alpha If Deleted
Supervisor Sample				
25. confident	98	0.84	0.44	0.89
26. likes	98	0.88	0.38	0.77
27. trust	98	0.92	0.33	0.71
Clinician Sample				
24. confident	96	0.83	0.54	0.83
25. likes	96	0.87	0.49	0.71
26. trust	96	0.88	0.45	0.70
Client Sample				
19. confident	77	0.88	0.78	0.81
20. likes	77	0.87	0.78	0.80
18. trust	77	0.89	0.79	0.73
Factor 4: Suicide Assessment				
Item	Obs	Item-Domain Correlation	Inter-Item Covariance	Alpha If Deleted
Supervisor Sample				
9. socrat	NA	NA	NA	NA
10. facilitate	NA	NA	NA	NA
18. Th feedback	NA	NA	NA	NA
15. struggle	78	0.12	0.10	0.19
18. th feedback	NA	NA	NA	NA
28.b. today suic	78	0.24	0.18	0.74
29.a. ever use suic	NA	NA	NA	NA
29.b. today use suic	78	0.22	0.17	0.61
Clinician Sample				
8. socrat	NA	NA	NA	NA
9. facilitate	85	0.60	0.24	0.20
17.th feedback	85	0.31	0.21	0.25
14. struggle	NA	NA	NA	NA
17. th feedback	NA	NA	NA	NA
27.b. today suic	85	0.26	0.19	0.39
28.a. ever use suic	85	0.64	0.23	0.30
28.b. today use suic	85	0.27	0.20	0.53
Client Sample				
5. socrat	56	0.91	0.08	0.40
6. facilitate	56	0.93	0.07	0.37
11. Th feedback	NA	NA	NA	NA
10. struggle	56	0.98	0.02	0.35
12. th feedback	56	0.57	0.24	0.59
21.b. today suic	56	0.43	0.24	0.40
Question not included in client survey	NA*	NA*	NA*	NA*
Question not included in client survey	NA*	NA*	NA*	NA*

TABLE J.3 (continued)				
Factor 5: Homework				
Item	Obs	Item-Domain Correlation	Inter-Item Covariance	Alpha If Deleted
Supervisor Sample				
19. assign	79	0.44	0.25	0.60
20. review instruc	79	0.46	0.25	0.64
21. address	79	0.15	0.13	0.61
22. solution	79	0.13	0.11	0.57
23. review hmwk	79	0.24	0.18	0.59
Clinician Sample				
18. assign	82	0.61	0.24	0.64
19. review instruc	82	0.51	0.25	0.61
20. address	82	0.24	0.18	0.64
21. solution	82	0.27	0.20	0.50
22. review hmwk	82	0.28	0.20	0.57
Client Sample				
13. assign	62	0.65	0.23	0.78
14. review instruc	62	0.68	0.22	0.81
Question not included in client survey	NA*	NA*	NA*	NA*
15. solution	62	0.69	0.21	0.66
16. review hmwk	62	0.66	0.22	0.77
NA = Question did not load onto this factor for the model indicated in the column header. NA* = Question was not included in the client version of the survey.				

TABLE J.4. Summary of Reliability Analysis by Measure Domain				
Factor	Statistical Test	Supervisor Sample	Clinician Sample	Client Sample
Structuring and conducting the session	Reliability (KR20)	0.88	0.78	0.77
	Average difficulty	0.65	0.58	0.85
	Average item-rest correlation	0.66	0.46	0.49
Psychoeducation and therapeutic techniques	Reliability (KR20)	0.89	0.83	0.77
	Average difficulty	0.56	0.50	0.80
	Average item-rest correlation	0.52	0.45	0.45
Therapeutic alliance	Reliability (Alpha)	0.85	0.82	0.84
	Inter-Item covariance	0.39	0.50	0.78
	Average item-rest correlation	0.73	0.67	0.72
Suicide assessment	Reliability (KR20)	0.69	0.58	0.64
	Average difficulty	0.19	0.41	0.76
	Average item-rest correlation	0.52	0.33	0.42
Homework	Reliability (KR20)	0.81	0.81	0.90
	Average difficulty	0.28	0.38	0.67
	Average item-rest correlation	0.60	0.59	0.76
NOTE: Domain difficulty represents the average item difficulty (percent facilities answering "Yes" to an item) across the domain and is only calculated for binary items. Inter-item covariance is the measure of the average covariance between the items and is only calculated for continuous items.				

APPENDIX K. ITEM-LEVEL INTER-RATER AGREEMENT

TABLE K.1. Item-Level Inter-Rater Agreement Between Supervisors and Clinicians				
Supervisor-Clinician	Observed Agreement	AC1	Significance Level	Confidence Interval
Did the therapist set an agenda?	81.82%	0.65	< 0.01	(0.48-0.82)
Did the therapist go over the agenda with the client?	81.33%	0.63	< 0.01	(0.45-0.81)
Did the therapist provide background on the treatment rationales and concepts during this session (i.e., why you are asking the client to do something or explaining why something is occurring within the session)?	73.24%	0.57	< 0.01	(0.38-0.77)
Did the therapist discuss or check-in on the client's treatment expectations (i.e., what will happen, how treatment will progress, expectations for improvement)?	66.67%	0.37	< 0.01	(0.14-0.60)
Did the therapist and his/her client mutually set or check-in on goals for treatment?	65.71%	0.43	< 0.01	(0.20-0.66)
Did the therapist identify salient problem areas related to the trauma? <ul style="list-style-type: none"> • Problem areas might include self-blame, other blame, power and control issues, beliefs impacted by the trauma (e.g., the world is a dangerous place), self-esteem, safety, trust, intimacy, and perception of danger. 	80.22%	0.70	< 0.01	(0.55-0.84)
Did the therapist use cognitive restructuring techniques (techniques to address cognitive issues such as negative thoughts, distortions, false beliefs or perceptions and replace them with accurate and more useful cognitions) to work on the identified problem areas?	64.71%	0.40	< 0.01	(0.19-0.61)
Did the therapist use a Socratic discussion method, that is, statements or questions designed for the client to examine their beliefs? <ul style="list-style-type: none"> • How do you know this? Can you give me an example? • What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? • How did you come to this conclusion? What evidence do you have to justify this? 	59.21%	0.31	0.01	(0.08-0.55)
Did the therapist facilitate the development of alternative hypotheses (i.e., alternative viewpoints or explanations) to problematic beliefs? Examples of alternative hypotheses to problematic thinking might include: 1) Distortion: People in authority can't be trusted. (More Helpful Thought: People in authority are individuals, and they don't all share the same strengths and weaknesses) 2) Distortion: Everyone is out to hurt me. I can't trust anyone. (More Helpful Thought: There are some dangerous people out there, but not everyone is out to harm you)	63.77%	0.39	< 0.01	(0.15-0.63)
Did the therapist identify areas of trauma related avoidance, where the trauma has shifted or restricted daily patterns of living (i.e., the trauma has influenced daily functioning)? For example, a client may avoid places with loud noises and lots of people.	60.49%	0.24	0.04	(0.01-0.46)
Did the therapist use techniques to systematically approach areas of trauma related avoidance, where the trauma has shifted daily patterns of living (i.e., the trauma has influenced daily functioning) from easier to more difficult situations? <ul style="list-style-type: none"> • For example, a person in a motor vehicle accident may be fearful of driving. An approach from easy to more difficult might look like: <ul style="list-style-type: none"> • Easy: Encouraging the client to ride in a car as a passenger for a short period of time. • Difficult: Encouraging the client to drive on street and then a freeway, etc. 	60.92%	0.29	0.01	(0.07-0.51)

TABLE K.1 (continued)				
Supervisor-Clinician	Observed Agreement	AC1	Significance Level	Confidence Interval
Did the therapist use any of the following techniques to deal with trauma related avoidance?				
a. Ask the client to imagine their traumatic experience for longer than 10 minutes	87.36%	0.83	< 0.01	(0.73-0.94)
b. Ask the client to write about their traumatic experience	72.62%	0.59	< 0.01	(0.42-0.77)
c. Socratic discussion method (i.e., "How do you know this? Can you give me an example?")	46.99%	-0.06	0.59	(0.00-0.16)
d. Real world experiments like visiting a place related to the traumatic experience with the client for longer than 10 minutes	83.53%	0.79	< 0.01	(0.68-0.91)
Did the therapist discuss and process the details of the client's recounting of the trauma, including the emotions surrounding the event?	62.50%	0.27	0.01	(0.06-0.48)
Did the therapist struggle to manage time for any of the reasons below:				
<ul style="list-style-type: none"> • Client talked incessantly or tangentially • Client had trouble keeping on task • Session time was abbreviated • You had trouble keeping the Client on task 	75.64%	0.63	< 0.01	(0.45-0.80)
Was the therapist directive (i.e., followed the agenda or guided the client to relevant discussion) during this session?	89.29%	0.87	< 0.01	(0.78-0.96)
Did the therapist ask his/her client for feedback or input on their treatment (i.e., "how is this working?"; "Are we working on things that you think are important?")? This would <i>exclude</i> progress monitoring.	72.37%	0.50	< 0.01	(0.30-0.71)
Did the therapist ask your client for feedback on himself/herself?	69.33%	0.48	< 0.01	(0.27-0.69)
Did the therapist assign his/her client homework or practice assignments (to be completed by the <i>next session</i>) to deal with issues surrounding PTSD symptoms (i.e., avoidance, thought monitoring, problematic beliefs, anxiety) or issues related to the trauma?	69.66%	0.41	< 0.01	(0.21-0.60)
Did the therapist review the assignment instructions and verify the client has a thorough understanding of the homework for the <i>next session</i> ?	75.00%	0.50	< 0.01	(0.30-0.70)
Did the therapist address difficulties or barriers related to completing of homework from the <i>previous session</i> ?	74.39%	0.60	< 0.01	(0.42-0.78)
Did the therapist work with his/her client to come up with solutions to difficulties, barriers, or issues in completing the homework from the <i>previous session</i> ?	70.73%	0.55	< 0.01	(0.36-0.74)
Did the therapist review and discuss the client's homework assigned during the <i>previous session</i> ?	69.23%	0.47	< 0.01	(0.27-0.68)
When reviewing the homework assigned from the <i>previous session</i> , did the therapist encourage the client or provide him/her with constructive feedback?	75.00%	0.58	< 0.01	(0.39-0.78)
I am confident in the therapist's ability to help this client.	28.42%	0.85	< 0.01	(0.80-0.89)
I believe this client likes the therapist.	47.37%	0.94	< 0.01	(0.91-0.96)
This client and the therapist have built mutual trust.	37.89%	0.92	< 0.01	(0.90-0.95)
a. To your knowledge, has the therapist ever conducted a suicide risk assessment with this client?	84.27%	0.80	< 0.01	(0.69-0.91)
b. Did the therapist conduct a suicide risk assessment during this session?	82.95%	0.72	< 0.01	(0.58-0.87)
a. To your knowledge, has the therapist ever used information from your suicide risk assessment to influence treatment or monitor progress with this client?	67.86%	0.46	< 0.01	(0.26-0.67)
b. Did the therapist use information from the suicide risk assessment to influence treatment or monitor progress during this session?	76.47%	0.63	< 0.01	(0.47-0.80)
a. To your knowledge, has the therapist used any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change?	80.23%	0.64	< 0.01	(0.47-0.81)
b. Did the therapist use any valid standardized instruments (e.g., The Revised PTSD Checklist) or psychometric scales to monitor PTSD symptoms and assess change during this session?	90.80%	0.89	< 0.01	(0.81-0.97)

TABLE K.1 (continued)				
Supervisor-Clinician	Observed Agreement	AC1	Significance Level	Confidence Interval
a. To your knowledge has the therapist ever provided the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.)?	93.90%	0.94	< 0.01	(0.88-0.99)
b. Did the therapist provide the client with education on their symptoms (i.e., education on avoidance, flashbacks, etc.) during this session?	65.91%	0.35	< 0.01	(0.15-0.56)
a. To your knowledge, has the therapist ever provided the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma)? <ul style="list-style-type: none"> • For example, this might include education on the nature of acquaintance rape vs. stranger rape, how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault. 	68.06%	0.49	< 0.01	(0.28-0.71)
b. Did the therapist provide the client with specific education on the nature of the traumatic event (i.e., changes in viewpoint or perception or facts about the type of trauma) during this session? <ul style="list-style-type: none"> • For example, this might include education on the nature of acquaintance rape vs. stranger rape, or how sexual assault generally influences view points and beliefs, or how a perpetrator may groom their victim before an assault. 	58.33%	0.21	0.07	(0.00-0.43)
a. To your knowledge, has the therapist ever provided the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment)?	80.00%	0.74	< 0.01	(0.59-0.89)
b. Did the therapist provide the client with an outline or overview of the treatment process (i.e., what will happen over the course of treatment) <i>during this session</i> ?	71.43%	0.47	< 0.01	(0.27-0.66)

TABLE K.2. Item-Level Inter-Rater Agreement Between Supervisors and Clients				
Item	Observed Agreement	AC1	Significance Level	Confidence Interval
Did you and your therapist discuss an agenda or plan for your session?	56.92%	0.32	0.02	(0.06-0.58)
Did your therapist talk about or check-in on your expectations of how therapy will go?	62.71%	0.37	0.01	(0.11-0.63)
Did your therapist work with you to set goals you both agreed on?	67.74%	0.56	< 0.01	(0.35-0.77)
Did your therapist help you become aware of or realize feelings, views or thoughts in your life that have been influenced by your traumatic experience? <ul style="list-style-type: none"> These might include feelings, views, or thoughts about being safe in the world, the presence of danger, trust, and self-esteem. 	81.33%	0.75	< 0.01	(0.61-0.89)
Did your therapist ask you several direct questions to <i>make you think critically about or examine</i> your thoughts, feelings, or beliefs? For example, your therapist might ask: <ul style="list-style-type: none"> How do you know this? Can you give me an example? What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? How did you come to this conclusion? What evidence do you have to justify this? 	80.33%	0.75	< 0.01	(0.59-0.90)
Did your therapist offer other ways of thinking about your issues (e.g., problem areas or areas you want to work on) related to the trauma? For example: <ul style="list-style-type: none"> Thought: "I can't trust anyone." Thought suggested by therapist: "Some people can't be trusted, but there are other people who are trustworthy." 	81.03%	0.76	< 0.01	(0.60-0.91)
Did you and your therapist discuss people, events, or places you now avoid or stay away from because of your traumatic experience? For example, someone in a car accident might avoid driving on the freeway.	72.13%	0.56	< 0.01	(0.35-0.78)
Did your therapist do <i>any</i> of the following things to help you deal with fear, anxiety or things you now avoid because of your trauma? <ol style="list-style-type: none"> Ask you to imagine or retell your traumatic experience for longer than 10 minutes Ask you to write about your traumatic experience Ask you questions to make you think critically about or examine your thoughts, feelings, or beliefs related to your fear, anxiety, and avoidance of things (i.e., "How do you know this? Can you give me an example?") Ask you to do real world experiments like visiting a place related to the traumatic experience for longer than 10 minutes 	68.06%	0.45	< 0.01	(0.23-0.67)
After you described your traumatic experience, did you and your therapist discuss the details of what happened to you, how it impacted your life, or your emotions about the event?	52.78%	0.15	0.25	(0.00-0.41)
Did your therapist make good use of your session time today?	15.38%	-0.67	< 0.01	(0.00-0.00)
Did your therapist ask for your opinion on how your treatment is going?	67.21%	0.48	< 0.01	(0.24-0.72)
Did your therapist ask for feedback on how she/he is doing in helping you recover from your PTSD?	59.26%	0.20	0.16	(0.00-0.47)
Did your therapist assign homework or practice <i>assignments (to be completed by the next session)</i> to work on your PTSD symptoms or problem areas?	64.79%	0.30	0.01	(0.08-0.53)
Did your therapist make sure you understood how to complete your homework for the <i>next session</i> ?	63.49%	0.30	0.02	(0.05-0.54)
If you had problems completing your <i>previously</i> assigned homework, did your therapist work with you to come up with solutions to these problems?	39.34%	-0.19	0.15	(0.00-0.07)
Did your therapist review and discuss your homework from the <i>previous session</i> ?	50.00%	0.01	0.92	(0.00-0.28)

TABLE K.2 (continued)				
Item	Observed Agreement	AC1	Significance Level	Confidence Interval
When reviewing the homework from the <i>previous session</i> , did your therapist encourage or provide you with constructive feedback?	44.44%	-0.07	0.61	(0.00-0.22)
My therapist and I have built mutual trust.	28.95%	0.90	< 0.01	(0.86-0.95)
I am confident in my therapist's ability to help me.	35.53%	0.88	< 0.01	(0.81-0.94)
I believe my therapist likes me as a person.	36.84%	0.89	< 0.01	(0.84-0.94)
a. Has your therapist ever asked you if you have had thoughts about committing suicide?	86.11%	0.84	< 0.01	(0.73-0.95)
b. <i>During this session</i> , did your therapist ask you if you had thoughts about committing suicide?	76.81%	0.58	< 0.01	(0.38-0.78)
a. Has your therapist ever asked you to answer questions about your PTSD symptoms? This might include completing a form before or after therapy.	50.00%	0.04	0.76	(0.00-0.31)
b. <i>During this session</i> , did your therapist ask you about your PTSD symptoms? This might include completing a form or survey before or after therapy.	37.88%	-0.21	0.11	(0.00-0.05)
a. Has your therapist ever provided information about PTSD and PTSD symptoms?	83.33%	0.80	< 0.01	(0.66-0.93)
b. <i>During this session</i> , did your therapist provide information about PTSD and PTSD symptoms?	56.06%	0.16	0.21	(0.00-0.42)
a. Has your therapist ever provided with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? • For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your viewpoints and beliefs	77.78%	0.68	< 0.01	(0.49-0.88)
b. <i>During this session</i> , did your therapist ever provide you with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? • For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your view points and beliefs	59.38%	0.20	0.11	(0.00-0.45)
a. Has your therapist ever explained how your particular treatment will work?	79.31%	0.75	< 0.01	(0.59-0.90)
b. <i>During this session</i> , did your therapist explain how your particular treatment will work?	42.86%	-0.08	0.56	(0.00-0.20)
NOTE: AC1 values above 0.80 suggest high agreement; 0.61-0.80 substantial agreement, 0.41-0.60 moderate agreement, 0.21-0.40 fair agreement, and 0-0.20 slight agreement.				

TABLE K.3. Item-Level Inter-Rater Agreement Between Clinicians and Clients				
Item	Observed Agreement	AC1	Significance Level	Confidence Interval
Did you and your therapist discuss an agenda or plan for your session?	65.33%	0.48	< 0.01	(0.27-0.69)
Did your therapist talk about or check-in on your expectations of how therapy will go?	64.79%	0.44	< 0.01	(0.21-0.67)
Did your therapist work with you to set goals you both agreed on?	74.63%	0.66	< 0.01	(0.49-0.84)
Did your therapist help you become aware of or realize feelings, views or thoughts in your life that have been influenced by your traumatic experience? <ul style="list-style-type: none"> • These might include feelings, views, or thoughts about being safe in the world, the presence of danger, trust, and self-esteem. 	80.26%	0.74	< 0.01	(0.60-0.88)
Did your therapist ask you several direct questions <i>to make you think critically about or examine</i> your thoughts, feelings, or beliefs? For example, your therapist might ask: <ul style="list-style-type: none"> • How do you know this? Can you give me an example? • What are some other ways of viewing this? What are the pros and cons to your way of thinking about this? • How did you come to this conclusion? What evidence do you have to justify this? 	65.75%	0.50	< 0.01	(0.29-0.71)
Did your therapist offer other ways of thinking about your issues (e.g., problem areas or areas you want to work on) related to the trauma? For example: <ul style="list-style-type: none"> • Thought: "I can't trust anyone." • Thought suggested by therapist: "Some people can't be trusted, but there are other people who are trustworthy." 	57.75%	0.37	< 0.01	(0.13-0.61)
Did you and your therapist discuss people, events, or places you now avoid or stay away from because of your traumatic experience? For example, someone in a car accident might avoid driving on the freeway.	58.33%	0.30	0.02	(0.06-0.55)
Did your therapist do any of the following things to help you deal with fear, anxiety or things you now avoid because of your trauma? a) Ask you to imagine or retell your traumatic experience <i>for longer than 10 minutes</i> b) Ask you to write about your traumatic experience c) Ask you questions to make you think critically about or examine your thoughts, feelings, or beliefs related to your fear, anxiety, and avoidance of things (i.e., "How do you know this? Can you give me an example?") d) Ask you to do real world experiments like visiting a place related to the traumatic experience <i>for longer than 10 minutes</i>	62.50%	0.32	0.01	(0.08-0.56)
After you described your traumatic experience, did you and your therapist discuss the details of what happened to you, how it impacted your life, or your emotions about the event?	56.76%	0.21	0.10	(0.00-0.45)
Did your therapist make good use of your session time today?	32.00%	-0.25	0.06	(0.00-0.01)
Did your therapist ask for your opinion on how your treatment is going?	66.67%	0.48	< 0.01	(0.25-0.70)
Did your therapist ask for feedback on how she/he is doing in helping you recover from your PTSD?	51.61%	0.04	0.75	(0.00-0.30)
Did your therapist assign homework or practice assignments <i>(to be completed by the next session)</i> to work on your PTSD symptoms or problem areas?	66.67%	0.38	< 0.01	(0.15-0.61)
Did your therapist make sure you understood how to complete your homework for the <i>next session</i> ?	69.70%	0.44	< 0.01	(0.21-0.67)

TABLE K.3 (continued)				
Item	Observed Agreement	AC1	Significance Level	Confidence Interval
If you had problems completing your previously assigned homework, did your therapist work with you to come up with solutions to these problems?	36.51%	-0.27	0.03	(0.00-0.00)
Did your therapist review and discuss your homework from the <i>previous session</i> ?	53.13%	0.06	0.62	(0.00-0.31)
When reviewing the homework from the <i>previous session</i> , did your therapist encourage or provide you with constructive feedback?	53.33%	0.07	0.61	(0.00-0.32)
My therapist and I have built mutual trust.	26.32%	0.88	< 0.01	(0.84-0.92)
I am confident in my therapist's ability to help me.	13.16%	0.78	< 0.01	(0.72-0.84)
I believe my therapist likes me as a person.	26.32%	0.86	< 0.01	(0.81-0.91)
a. Has your therapist ever asked you if you have had thoughts about committing suicide?	85.14%	0.81	< 0.01	(0.70-0.93)
b. <i>During this session</i> , did your therapist ask you if you had thoughts about committing suicide?	68.12%	0.41	< 0.01	(0.18-0.64)
a. Has your therapist ever asked you to answer questions about your PTSD symptoms? This might include completing a form before or after therapy.	45.59%	-0.04	0.78	(0.00-0.23)
b. <i>During this session</i> , did your therapist ask you about your PTSD symptoms? This might include completing a form or survey before or after therapy.	40.85%	-0.14	0.26	(0.00-0.11)
a. Has your therapist ever provided information about PTSD and PTSD symptoms?	82.35%	0.77	< 0.01	(0.63-0.91)
b. <i>During this session</i> , did your therapist provide information about PTSD and PTSD symptoms?	69.57%	0.41	< 0.01	(0.19-0.63)
a. Has your therapist ever provided with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? • For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your viewpoints and beliefs	60.00%	0.35	0.01	(0.10-0.60)
b. <i>During this session</i> , did your therapist ever provide you with specific education on the nature of the traumatic event (i.e., facts about the type of trauma)? • For example, this might include education on the nature of sexual assault, or how sexual assault generally influences your view points and beliefs	50.00%	0.00	0.99	(0.00-0.25)
a. Has your therapist ever explained how your particular treatment will work?	87.32%	0.85	< 0.01	(0.75-0.96)
b. <i>During this session</i> , did your therapist explain how your particular treatment will work?	51.47%	0.10	0.45	(0.00-0.36)
NOTE: AC1 values above 0.80 suggest high agreement; 0.61-0.80 substantial agreement, 0.41-0.60 moderate agreement, 0.21-0.40 fair agreement, and 0-0.20 slight agreement.				

APPENDIX L. COMPARISON BETWEEN BAYESIAN AND WEIGHTED LEAST SQUARES AND MEAN VARIANCE (WLSMV) ADJUSTED ESTIMATORS

**TABLE L.1. Comparison Between Bayesian and WLSMV Adjusted Estimators:
Clinical Confirmatory Factor-Analytic Models**

Item and Item Number	Bayesian Estimates						WLSMV Estimates			
	Estimate	Posterior S.D.	One-Tailed p	95% C.I.		Signif.	Estimate	S.E.	Est/ S.E.	p-value
				Lower 2.5%	Upper 2.5%					
F1										
1. AGENDA	0.98	0.02	0.00	0.91	0.99	*	0.71	0.08	8.43	0.00
2. REVIEW AGENDA	0.98	0.02	0.00	0.93	0.99	*	0.73	0.08	9.00	0.00
4. EXPECTATIONS	0.60	0.11	0.00	0.35	0.78	*	0.78	0.08	9.39	0.00
5. GOALS	0.41	0.14	0.00	0.11	0.65	*	0.45	0.13	3.62	0.00
15. DIRECTIVE	0.79	0.11	0.00	0.50	0.93	*	0.57	0.15	3.92	0.00
16. TX FEEDBACK	0.49	0.13	0.00	0.22	0.71	*	0.69	0.09	7.33	0.00
29.b. TODAY INSTRUC	0.83	0.11	0.00	0.51	0.95	*	0.69	0.15	4.77	0.00
32.a. EVER OUTLINE	0.45	0.16	0.01	0.10	0.69	*	0.47	0.19	2.49	0.01
32.b. TODAY OUTLINE	0.83	0.07	0.00	0.67	0.92	*	1.05	0.08	13.59	0.00
F2										
12.a. IMAGINE	0.69	0.12	0.00	0.42	0.87	*	0.66	0.11	5.92	0.00
12.b. WRITE	0.88	0.07	0.00	0.68	0.96	*	0.92	0.07	14.05	0.00
12.c. OTHER SOCRAT	0.93	0.06	0.00	0.76	0.98	*	0.79	0.06	12.33	0.00
12.d. REAL	0.74	0.12	0.00	0.43	0.90	*	0.67	0.14	4.74	0.00
3. BACKGROUND	0.68	0.10	0.00	0.43	0.83	*	0.86	0.07	12.05	0.00
6. IDENTIFY	0.65	0.11	0.00	0.39	0.82	*	0.71	0.11	6.54	0.00
7. COG RESTRUC	0.58	0.12	0.00	0.31	0.77	*	0.57	0.10	5.95	0.00
8. SOCRAT	0.79	0.09	0.00	0.56	0.92	*	0.75	0.07	11.23	0.00
10. OTHER IDENTIFY	0.52	0.12	0.00	0.25	0.71	*	0.60	0.09	6.51	0.00
11. TECHNIQUES	0.52	0.12	0.00	0.24	0.71	*	0.50	0.11	4.69	0.00
13. DISCUSS	0.64	0.10	0.00	0.40	0.80	*	0.60	0.09	6.55	0.00
30.a. EVER SYMP EDU	0.90	0.06	0.00	0.72	0.96	*	0.86	0.16	5.43	0.00
30.b. TODAY SYMP EDU	0.65	0.11	0.00	0.40	0.81	*	0.75	0.07	10.98	0.00
31.b. TODAY TRAUMA ED	0.49	0.13	0.00	0.21	0.71	*	0.56	0.11	5.31	0.00
12. OVERALL TECHNIQUES	0.98	0.02	0.00	0.93	0.99	*	0.80	0.06	13.92	0.00
F3										
24. CONFIDENT	0.65	0.07	0.00	0.49	0.78	*	1.06	0.28	3.74	0.00
25. LIKES	0.83	0.06	0.00	0.69	0.92	*	0.46	0.13	3.67	0.00
26. TRUST	0.87	0.05	0.00	0.75	0.96	*	0.55	0.15	3.67	0.00
F4										
9. FACILITATE	0.48	0.14	0.00	0.17	0.71	*	0.62	0.15	4.28	0.00
17. TH FEEDBACK	0.68	0.13	0.00	0.39	0.90	*	0.77	0.14	5.69	0.00
27.b. TODAY SUIC	0.76	0.11	0.00	0.50	0.93	*	0.67	0.15	4.38	0.00
28.a. EVER USE SUIC	0.58	0.14	0.00	0.26	0.80	*	0.44	0.14	3.11	0.00
28.b. TODAY USE SUIC	0.92	0.07	0.00	0.70	0.97	*	0.77	0.12	6.36	0.00
F5										
18. ASSIGN	0.96	0.03	0.00	0.89	0.99	*	0.94	0.05	20.44	0.00
19. REVIEW INST	0.94	0.04	0.00	0.83	0.98	*	1.01	0.04	23.35	0.00
20. ADDRESS	0.91	0.06	0.00	0.76	0.97	*	0.97	0.05	20.12	0.00
21. SOLUTION	0.80	0.08	0.00	0.58	0.91	*	0.77	0.08	9.79	0.00
22. REVIEW HMWK	0.83	0.08	0.00	0.61	0.94	*	0.79	0.06	12.74	0.00
F2										
F1	0.49	0.10	0.00	0.26	0.67	*	0.66	0.07	9.40	0.00
F3										
F1	-0.01	0.12	0.46	-0.26	0.22		0.14	0.12	1.12	0.27
F2	0.12	0.12	0.17	-0.12	0.34		0.21	0.11	1.95	0.05

TABLE L.1. (continued)										
Item and Item Number	Bayesian Estimates						WLSMV Estimates			
	Estimate	Posterior S.D.	One-Tailed p	95% C.I.		Signif.	Estimate	S.E.	Est/S.E.	p-value
				Lower 2.5%	Upper 2.5%					
F4										
F1	0.37	0.13	0.00	0.09	0.60	*	0.50	0.11	4.52	0.00
F2	0.37	0.12	0.00	0.11	0.59	*	0.49	0.11	4.62	0.00
F3	-0.14	0.13	0.16	-0.39	0.12		0.01	0.13	0.05	0.96
F5										
F1	0.37	0.11	0.00	0.14	0.57	*	0.40	0.11	3.67	0.00
F2	0.51	0.10	0.00	0.29	0.68	*	0.52	0.08	6.28	0.00
F3	0.24	0.12	0.03	-0.01	0.47		0.30	0.11	2.61	0.01
F4	0.19	0.14	0.10	-0.10	0.46		0.19	0.13	1.43	0.15

NOTE: In Bayesian models "significance" (*) indicates that the true factor loading will lie within the estimated credibility interval with the 95% posterior probability. A blank cell indicates the factor is not statistically significant.

**TABLE L.2. Comparison Between Bayesian and WLSMV Adjusted Estimators:
Supervisor Confirmatory Factor-Analytic Models**

Item and Item Number	Bayesian Estimates						WLSMV Estimates			
	Estimate	Posterior S.D.	One-Tailed p	95% C.I.		Signif.	Estimate	S.E.	Est/S.E.	p-value
				Lower 2.5%	Upper 2.5%					
F1										
2. AGENDA	0.96	0.02	0.00	0.90	0.99	*	0.96	0.06	16.71	0.00
3. REVIEW AGENDA	0.97	0.02	0.00	0.93	0.99	*	1.01	0.03	29.54	0.00
5. EXPECTATIONS	0.85	0.07	0.00	0.69	0.94	*	0.89	0.06	15.68	0.00
6. GOALS	0.75	0.09	0.00	0.53	0.88	*	0.80	0.08	10.50	0.00
16. DIRECTIVE	0.76	0.12	0.00	0.45	0.91	*	0.64	0.13	5.04	0.00
17. TX FEEDBACK	0.94	0.04	0.00	0.84	0.98	*	0.94	0.05	20.75	0.00
33.b. TODAY OUTLINE	0.80	0.08	0.00	0.60	0.91	*	0.81	0.08	10.48	0.00
F2										
13.a. IMAGINE	0.73	0.13	0.00	0.42	0.90	*	0.57	0.12	4.66	0.00
13.b. WRITE	0.76	0.09	0.00	0.53	0.89	*	0.77	0.08	9.40	0.00
13.c. OTHER SOCRAT	0.91	0.04	0.00	0.81	0.96	*	0.87	0.06	15.54	0.00
13.d. REAL	0.70	0.12	0.00	0.42	0.88	*	0.76	0.09	8.38	0.00
4. BACKGROUND	0.81	0.07	0.00	0.62	0.91	*	0.83	0.07	12.38	0.00
7. IDENTIFY	0.83	0.07	0.00	0.64	0.92	*	0.80	0.08	9.66	0.00
8. COG RESTRUC	0.56	0.12	0.00	0.27	0.75	*	0.60	0.11	5.24	0.00
9. SOCRAT	0.47	0.14	0.00	0.14	0.70	*	0.46	0.13	3.44	0.00
11. OTHER IDENTIFY	0.85	0.06	0.00	0.70	0.94	*	0.90	0.06	16.09	0.00
12. TECHNIQUES	0.79	0.08	0.00	0.60	0.90	*	0.87	0.05	16.48	0.00
14. DISCUSS	0.80	0.07	0.00	0.63	0.90	*	0.80	0.07	11.04	0.00
19. ASSIGN	0.59	0.10	0.00	0.41	0.78	*	0.69	0.12	5.91	0.00
20. REVIEW INSTRUC	0.52	0.10	0.00	0.32	0.72	*	0.61	0.12	5.14	0.00
31.b. TODAY SYMP EDU	0.85	0.06	0.00	0.71	0.93	*	0.82	0.07	12.07	0.00
32.a. EVER TRAUMA ED	0.83	0.08	0.00	0.61	0.93	*	0.77	0.09	8.27	0.00
31.b. TODAY TRAUMA ED	0.93	0.04	0.00	0.83	0.98	*	0.91	0.04	20.95	0.00
12. OVERALL TECHNIQUES	0.94	0.04	0.00	0.84	0.98	*	0.92	0.04	21.04	0.00
F3										
25. CONFIDENT	0.66	0.07	0.00	0.52	0.77	*	0.53	0.07	7.34	0.00
26. LIKES	0.85	0.04	0.00	0.76	0.93	*	0.62	0.05	12.74	0.00
27. TRUST	0.95	0.03	0.00	0.87	0.99	*	0.72	0.05	13.42	0.00
F4										
15. STRUGGLE	0.56	0.19	0.01	0.12	0.83	*	0.45	0.21	2.10	0.04
28.b. TODAY SUIC	0.96	0.03	0.00	0.88	0.99	*	0.97	0.11	8.75	0.00
29.b. TODAY USE SUIC	0.96	0.02	0.00	0.89	0.98	*	0.99	0.10	9.48	0.00
F5										
19. ASSIGN	0.75	0.07	0.00	0.60	0.89	*	0.69	0.10	6.97	0.00
20. REVIEW INSTRUC	0.80	0.08	0.00	0.63	0.94	*	0.75	0.12	6.18	0.00
21. ADDRESS	0.94	0.05	0.00	0.80	0.98	*	1.02	0.03	35.56	0.00
22. SOLUTION	0.84	0.11	0.00	0.53	0.95	*	0.93	0.05	19.38	0.00
23. REVIEW HMWK	0.94	0.03	0.00	0.85	0.98	*	0.89	0.06	16.01	0.00
F2										
F1	0.68	0.08	0.00	0.51	0.81	*	0.72	0.08	9.53	0.00
F3										
F1	0.09	0.12	0.23	-0.14	0.32		0.08	0.10	0.79	0.43
F2	-0.02	0.11	0.43	-0.24	0.20		-0.03	0.10	-0.27	0.79
F4										
F1	0.27	0.15	0.05	-0.04	0.55		0.14	0.16	0.91	0.36
F2	0.23	0.14	0.05	-0.04	0.49		0.24	0.15	1.57	0.12
F3	-0.22	0.12	0.04	-0.44	0.03		-0.19	0.13	-1.50	0.13
F5										
F1	0.11	0.15	0.24	-0.19	0.39		0.04	0.13	0.32	0.75
F2	0.02	0.14	0.46	-0.25	0.27		0.03	0.15	0.21	0.83
F3	-0.09	0.12	0.23	-0.32	0.15		-0.03	0.12	-0.24	0.81
F4	0.62	0.13	0.00	0.32	0.83	*	0.45	0.13	3.44	0.00

NOTE: In Bayesian models "significance" (*) indicates that the true factor loading will lie within the estimated credibility interval with the 95% posterior probability. A blank cell indicates the factor is not statistically significant.

**TABLE L.3. Comparison Between Bayesian and WLSMV Adjusted Estimators:
Client Confirmatory Factor-Analytic Models**

Item and Item Number	Bayesian Estimates						WLSMV Estimates			
	Estimate	Posterior S.D.	One-Tailed p	95% C.I.		Signif.	Estimate	S.E.	Est/S.E.	p-value
				Lower 2.5%	Upper 2.5%					
F1										
1. AGENDA	0.78	0.13	0.00	0.44	0.93	*	0.71	0.10	6.90	0.00
2. EXPECTATIONS	0.90	0.06	0.00	0.72	0.96	*	0.95	0.07	13.95	0.00
3. GOALS	0.81	0.14	0.00	0.42	0.95	*	0.82	0.10	8.14	0.00
11. TX FEEDBACK	0.78	0.10	0.00	0.51	0.92	*	0.88	0.09	9.92	0.00
22.b. TODAY INSTR	0.68	0.13	0.00	0.38	0.87	*	0.70	0.11	6.52	0.00
25.a. EVER OUTLINE	0.79	0.18	0.00	0.29	0.97	*	0.69	0.06	11.31	0.00
25.b. TODAY OUTLI	0.87	0.08	0.00	0.66	0.96	*	0.82	0.09	9.14	0.00
F2										
4. IDENTIFY	0.64	0.17	0.00	0.24	0.88	*	0.64	0.17	3.70	0.00
7. OTHER IDENTIFY	0.45	0.17	0.02	0.03	0.72	*	0.41	0.16	2.60	0.01
9. DISCUSS	0.93	0.06	0.00	0.77	0.98	*	0.91	0.08	11.95	0.00
23.a. EVER SYMP EDU	0.68	0.14	0.00	0.34	0.87	*	0.74	0.12	6.23	0.00
23.b. TODAY SYMP EDU	0.93	0.05	0.00	0.79	0.98	*	0.89	0.06	15.91	0.00
24.a. EVER TRAUMA ED	0.83	0.09	0.00	0.59	0.94	*	0.85	0.08	11.34	0.00
24.b. TODAY TRAUMA ED	0.94	0.05	0.00	0.80	0.99	*	0.91	0.06	16.20	0.00
8. OVERALL TECHNIQUES	0.72	0.11	0.00	0.45	0.88	*	0.80	0.09	9.01	0.00
F3										
19. CONFIDENT	0.90	0.05	0.00	0.78	0.97	*	0.87	0.09	9.92	0.00
20. LIKES	0.78	0.06	0.00	0.63	0.87	*	0.81	0.11	7.70	0.00
18. TRUST	0.77	0.06	0.00	0.63	0.87	*	0.77	0.13	6.02	0.00
F4										
5. SOCRAT	0.58	0.19	0.01	0.15	0.87	*	0.64	0.16	4.08	0.00
6. FACILITATE	0.85	0.12	0.00	0.51	0.98	*	0.81	0.10	7.92	0.00
10. STRUGGLE	0.79	0.17	0.00	0.29	0.93	*	0.64	0.10	6.13	0.00
12. TH FEEDBACK	0.92	0.07	0.00	0.73	0.98	*	0.83	0.07	11.28	0.00
21.b. TODAY SUIC	0.63	0.15	0.00	0.29	0.87	*	0.51	0.10	5.37	0.00
F5										
13. ASSIGN	0.97	0.03	0.00	0.89	0.99	*	0.83	0.07	11.80	0.00
14. REVIEW INSTRUC	0.97	0.02	0.00	0.91	0.99	*	0.91	0.06	14.97	0.00
15. SOLUTION	0.79	0.09	0.00	0.57	0.91	*	0.94	0.06	15.83	0.00
16. REVIEW HMWK	0.88	0.07	0.00	0.69	0.96	*	0.97	0.05	20.37	0.00
F2										
F1	0.70	0.10	0.00	0.46	0.86	*	0.76	0.11	7.14	0.00
F3										
F1	0.13	0.15	0.19	-0.16	0.41		0.22	0.11	1.96	0.05
F2	0.28	0.14	0.03	0.00	0.53		0.29	0.11	2.53	0.01
F4										
F1	0.66	0.14	0.00	0.34	0.88	*	0.90	0.10	8.87	0.00
F2	0.74	0.10	0.00	0.51	0.91	*	0.86	0.06	14.08	0.00
F3	0.43	0.14	0.00	0.14	0.66	*	0.64	0.11	6.02	0.00
F5										
F1	0.55	0.14	0.00	0.25	0.77	*	0.72	0.10	6.86	0.00
F2	0.28	0.14	0.03	-0.02	0.54		0.39	0.14	2.93	0.00
F3	0.27	0.14	0.03	-0.02	0.54		0.27	0.13	1.98	0.05
F4	0.43	0.16	0.01	0.09	0.73	*	0.55	0.12	4.64	0.00
NOTE: In Bayesian models "significance" (*) indicates that the true factor loading will lie within the estimated credibility interval with the 95% posterior probability. A blank cell indicates the factor is not statistically significant.										

APPENDIX M. ACI INDEX

ACI Index

An analysis of the marginal Kappa distributions suggests that the lower values of Kappa may be attributed to the so-called “Kappa paradoxes” (Feinstein and Cicchetti, 1990), which occur when raters yield a high percent positive (responses in which both raters give a rating of “yes”) or negative agreement (both raters give a rating of “no”). Simply put, Kappa tends to yield a low value when the raters show high agreement, which is counterintuitive since one would expect a higher reliability when two qualified raters reach high agreement in observed ratings. Gwet’s adjusted chance-corrected AC1 index (Gwet 2008) was developed specifically to overcome these weaknesses of the Kappa statistic.

Gwet's agreement co-efficient can be used in more contexts than Kappa because it does not depend upon the assumption of independence between raters. The AC1 is based upon the more realistic assumption that only a portion of the observed ratings will potentially lead to agreement by chance.⁸ Gwet (2008) indicates that a reasonable value for chance-agreement probability should not exceed 0.5, whereas chance-agreement probability for Cohen’s (1960) Kappa can be any value between 0 and 1. For instance, if the raters agree 90 percent of time, Gwet’s AC1 would assume that chance-agreement should be at most 50 percent, whereas Kappa would calculate the chance-agreement at 81 percent on the positives and 1 percent on the negatives for a total of 82 percent.

This limit of chance-agreement prevents Gwet’s AC1 statistics from the form of erratic behavior seen in Kappa. Another beneficial property of AC1 is that while Cohen’s Kappa penalizes raters who produce similar ratings or marginal distributions, such penalization does not exist in AC1; on the contrary, raters with homogenous marginal distributions are rewarded by AC1.

⁸ To better understand what the AC1 for 2 raters conceptually represents, imagine that all subjects who are classified into identical categories by pure chance are identified and removed from the population of subjects. This creates a new trimmed population where agreement by chance would be impossible. The AC1 co-efficient is the relative number of subjects in the trimmed subject population upon which the raters agreed (Gwet 2008).

To obtain a printed copy of this report, send the full report title and your mailing information to:

U.S. Department of Health and Human Services
Office of Disability, Aging and Long-Term Care Policy
Room 424E, H.H. Humphrey Building
200 Independence Avenue, S.W.
Washington, D.C. 20201
FAX: 202-401-7733

NOTE: All requests must be in writing.

RETURN TO:

Office of Disability, Aging and Long-Term Care Policy (DALTCP) Home
<http://aspe.hhs.gov/office-disability-aging-and-long-term-care-policy-daltcp>

Assistant Secretary for Planning and Evaluation (ASPE) Home
<http://aspe.hhs.gov>

U.S. Department of Health and Human Services (HHS) Home
<http://www.hhs.gov>